

DOCKET NO.: 263365US0X PCT

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

IN RE APPLICATION OF: Mylene WEILL, et al.

SERIAL NO.: NEW U.S. PCT APPLICATION

FILED: HEREWITH

INTERNATIONAL APPLICATION NO.: PCT/FR03/01876

INTERNATIONAL FILING DATE: June 19, 2003

FOR: NOVEL ACETYLCHOLINESTERASE GENE RESPONSIBLE FOR INSECTICIDE  
RESISTANCE AND APPLICATIONS THEREOF**REQUEST FOR PRIORITY UNDER 35 U.S.C. 119  
AND THE INTERNATIONAL CONVENTION**Commissioner for Patents  
Alexandria, Virginia 22313

Sir:

In the matter of the above-identified application for patent, notice is hereby given that  
the applicant claims as priority:

<u>COUNTRY</u>	<u>APPLICATION NO</u>	<u>DAY/MONTH/YEAR</u>
France	02 07622	20 June 2002
France	02 13799	05 November 2002

Certified copies of the corresponding Convention application(s) were submitted to the International Bureau in PCT Application No. PCT/FR03/01876. Receipt of the certified copy(s) by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.

Respectfully submitted,  
OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.



Norman F. Oblon  
Attorney of Record  
Registration No. 24,618  
Surinder Sachar  
Registration No. 34,423

Corwin P. Umbach, Ph.D.  
Registration No. 40,211

Customer Number

**22850**

(703) 413-3000  
Fax No. (703) 413-2220  
(OSMMN 08/03)

BEST AVAILABLE COPY



PCT/R03/01876

REC'D 01 SEP 2003

WIPO PCT

# BREVET D'INVENTION

**CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION**

## COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 24 JUIN 2003

Pour le Directeur général de l'Institut  
national de la propriété industrielle  
Le Chef du Département des brevets

Martine PLANCHE

**DOCUMENT DE PRIORITÉ**

PRÉSENTÉ OU TRANSMIS  
CONFORMÉMENT À LA  
RÈGLE 17.1.a) OU b)

INSTITUT  
NATIONAL DE  
LA PROPRIÉTÉ  
INDUSTRIELLE

SIEGE  
26 bis, rue de Saint Petersburg  
75800 PARIS cedex 08  
Téléphone : 33 (0)1 53 04 53 04  
Télécopie : 33 (0)1 53 04 45 23  
www.inpi.fr



26 bis, rue de Saint Pétersbourg  
75800 Paris Cedex 08  
Téléphone : 01 53 04 53-04 Télécopie : 01 42 94 86 54

**BREVET D'INVENTION**  
**CERTIFICAT D'UTILITÉ**  
Code de la propriété intellectuelle - Livre VI



REQUÊTE EN DÉLIVRANCE 1/2

<b>REMISE DES PIÈCES</b> DATE <b>20 JUIN 2002</b> LIEU <b>75 INPI PARIS</b> N° D'ENREGISTREMENT <b>0207622</b> NATIONAL ATTRIBUÉ PAR L'INPI DATE DE DÉPÔT ATTRIBUÉE PAR L'INPI <b>20 JUIN 2002</b>		Cet imprimé est à remplir lisiblement à l'encre noire DB 540 W / 260G39	
<b>Vos références pour ce dossier</b> (facultatif) BLOcp644/80FR		<b>1 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE</b> À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE  CABINET ORES  6 avenue de Messine 75008 PARIS	
<b>Confirmation d'un dépôt par télécopie</b> <input type="checkbox"/> N° attribué par l'INPI à la télécopie			
<b>2 NATURE DE LA DEMANDE</b>		<b>Cochez l'une des 4 cases suivantes</b>	
Demande de brevet		<input checked="" type="checkbox"/>	
Demande de certificat d'utilité		<input type="checkbox"/>	
Demande divisionnaire		<input type="checkbox"/>	
Demande de brevet initiale		N° _____ Date ____/____/____	
ou demande de certificat d'utilité initiale		N° _____ Date ____/____/____	
Transformation d'une demande de brevet européen		<input type="checkbox"/> N° _____ Date ____/____/____	
<b>3 TITRE DE L'INVENTION</b> (200 caractères ou espaces maximum)			
NOUVEAU GENE DE L'ACETYLCHOLINESTERASE RESPONSABLE DE LA RESISTANCE AUX INSECTICIDES ET SES APPLICATIONS.			
<b>4 DÉCLARATION DE PRIORITÉ</b> OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE FRANÇAISE		Pays ou organisation _____ N° _____ Date ____/____/____ Pays ou organisation _____ N° _____ Date ____/____/____ Pays ou organisation _____ N° _____ Date ____/____/____ <input type="checkbox"/> S'il y a d'autres priorités, cochez la case et utilisez l'imprimé «Suite»	
<b>5 DEMANDEUR</b>		<input type="checkbox"/> S'il y a d'autres demandeurs, cochez la case et utilisez l'imprimé «Suite»	
Nom ou dénomination sociale		CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE-CNRS	
Prénoms			
Forme juridique		Etablissement public	
N° SIREN		. . . . .	
Code APE-NAF		. . . . .	
Adresse	Rue	3 rue Michel-Ange	
	Code postal et ville	75794 PARIS CEDEX 16	
Pays		FRANCE	
Nationalité		Française	
N° de téléphone (facultatif)			
N° de télécopie (facultatif)			
Adresse électronique (facultatif)			

REMISE DES PIÈCES DATE <b>25 JUIN 2002</b> LIEU <b>75 INPI PARIS</b> N° D'ENREGISTREMENT <b>0207622</b> NATIONAL ATTRIBUÉ PAR L'INPI		Réservé à l'INPI	DB 540 W / 260599
<b>Vos références pour ce dossier :</b> <i>(facultatif)</i>		BLOcp644/80FR	
<b>6 MANDATAIRE</b>			
Nom		ORES	
Prénom		Béatrice	
Cabinet ou Société		CABINET ORES	
N° de pouvoir permanent et/ou de lien contractuel			
Adresse	Rue	6 avenue de Messine	
	Code postal et ville	75008	PARIS
N° de téléphone <i>(facultatif)</i>		01.45.62.75.00.	
N° de télécopie <i>(facultatif)</i>		01.45.62.04.86.	
Adresse électronique <i>(facultatif)</i>		ores@cabinet-ores.com	
<b>7 INVENTEUR (S)</b>			
Les inventeurs sont les demandeurs		<input type="checkbox"/> Oui <input checked="" type="checkbox"/> Non Dans ce cas fournir une désignation d'inventeur(s) séparée	
<b>8 RAPPORT DE RECHERCHE</b>		Uniquement pour une demande de brevet (y compris division et transformation)	
Établissement immédiat ou établissement différé		<input checked="" type="checkbox"/> <input type="checkbox"/>	
Paiement échelonné de la redevance		Paiement en trois versements, uniquement pour les personnes physiques <input type="checkbox"/> Oui <input type="checkbox"/> Non	
<b>9 RÉDUCTION DU TAUX DES REDEVANCES</b>		Uniquement pour les personnes physiques <input type="checkbox"/> Requête pour la première fois pour cette invention <i>(joindre un avis de non-imposition)</i> <input type="checkbox"/> Requête antérieurement à ce dépôt <i>(joindre une copie de la décision d'admission pour cette invention ou indiquer sa référence)</i> :	
Si vous avez utilisé l'imprimé «Sulte», indiquez le nombre de pages jointes			
<b>10 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE</b> (Nom et qualité du signataire)  ORES Béatrice (n° 92-4046)		<b>VISA DE LA PRÉFECTURE OU DE L'INPI</b>  C. MARTIN	

La présente invention est relative à un nouveau gène de l'acétylcholinestérase responsable de la résistance aux insecticides, notamment chez les moustiques, aux produits de ce gène (ADNc, protéine) et à leurs applications, notamment pour le criblage de nouveaux insecticides et la détection génétique de la résistance aux organophosphorés et aux carbamates dans les populations de moustiques.

L'acétylcholinestérase (AChE, E.C. 3.1.1.7) est une enzyme essentielle qui hydrolyse l'acétylcholine dans les synapses, mettant ainsi fin aux transmissions cholinergiques au niveau des jonctions neuronales ou neuromusculaires. L'inhibition de l'AChE empêche la désactivation du signal synaptique, conduisant ainsi à une perte de contrôle de la transmission cholinergique. La biologie de l'acétylcholinestérase a été très étudiée chez les invertébrés, et en particulier les insectes, car cette enzyme est la cible des principales classes de pesticides utilisés, les organophosphorés et les carbamates. Cependant, l'utilisation massive de pesticides au cours des dernières décennies a provoqué l'émergence d'espèces résistantes. Parmi les mécanismes de résistance, la sélection de mutations rendant l'AChE insensible aux insecticides a été observée dans de nombreux cas (Pour une revue, voir Fournier et al., Comp. Biochem. Physiol., 1994, 108, 19-31).

Afin de déterminer avec précision, la nature de l'AChE cible des insecticides, ainsi que les mutations responsables de la résistance à ces derniers, les gènes codant pour des AChE (gènes *ace*) ont été isolés chez différentes espèces d'arthropodes (insectes et arachnides).

Le premier gène *ace* a été identifié chez la drosophile (*Drosophila melanogaster*), par génétique inverse (Hall et al., EMBO J., 1986, 5, 2949-2954). La preuve que ce gène était impliqué dans la résistance aux insecticides a été fournie par la mise en évidence de substitutions d'acides aminés dans l'AChE de drosophiles résistantes, conférant l'insensibilité aux insecticides cholinergiques (Mutéro et al., P.N.A.S., 1994, 91, 5922-5926). Les études chez *D. melanogaster* semblaient donc indiquer la présence d'un seul gène *ace* chez les insectes, codant pour l'AChE cible des insecticides cholinergiques.

Toutefois, à l'exception du gène *ace* d'un autre insecte, *Musca domestica* (Williamson et al., 1992, In *Multidisciplinary approaches to cholinesterase functions*, Eds Schafferman A. & Velan B., Plenum Press, New-York, pp 83-86 ;

Walsh et al., Biochem. J., 2001, 359, 175-181 ; Kozaki et al., Insect. Biochem. Mol. Biol., 2001, 31, 991-997), l'étude des gènes *ace* isolés chez d'autres insectes ou bien chez des arachnides, par homologie avec celui de la drosophile, indiquent qu'ils ne sont pas impliqués dans la résistance aux insecticides.

5 En effet, aucune mutation dans la séquence en acides aminés de l'AChE codée par le gène *ace* d'*Aphis gossypii*, de *Nephotettix cincticeps* et de

*Boophilus microplus* n'est observée entre les individus résistants et sensibles (Menozzi et al., Thèse de Doctorat de l'université Paul sabatier, Toulouse, 2000 ; Tomita et al., Insect Biochem. Mol. Biol., 200, 30, 325-333; Baxter et al., Insect Biochem. Mol. Biol., 1998, 28, 581-589; Hernandez et al., J. Med. Entomol., 1999, 36, 764-770), et une ségrégation indépendante est observée entre le gène *ace* de *Culex pipiens* et *C. tritaeniorhynchus* et la résistance aux insecticides (Malcolm et al., Insect. Mol. Biol., 1998, 7, 107-120; Mori et al., Insect Mol. Biol., 2001, 10, 197-203).

En ce qui concerne les autres gènes *ace* isolés chez d'autres insectes, leur rôle dans la résistance aux insecticides n'a pas été étudié (*Lucilia cuprina* : Chen et al., Insect. Biochem. Mol. Biol., 2001, 31, 805-816 ; *Schizaphis graminum* : Gao et al., Insect. Biochem. Mol. Biol., 2001, 31, 1095-1104) ou aucune forme d'AChE insensible aux insecticides n'a été décrite (*Aedes aegypti*, *Anopheles gambiae* et *Anopheles stephensi* : Anthony et al., FEBS letters, 1995, 368, 461-465 ; Malcolm et al., In *Molecular Insect Science*, Eds Hageborn et al., Plenum Press, New-York, pp 57-65).

Deux hypothèses ont été émises pour expliquer la différence dans la résistance aux insecticides, observée entre *Drosophila melanogaster* ou *Musca domestica* et les autres insectes ou les arachnides qui ont été étudiés : la présence d'un "gène modificateur" responsable de modifications post-transcriptionnelles ou post-traductionnelles de l'AChE, conduisant à des formes d'AChE possédant des activités catalytiques différentes, et la présence d'un deuxième gène *ace*.

Toutefois, aucune étude n'a permis de vérifier ces hypothèses et par conséquent de déterminer la nature du gène et celle de la cible (AChE) impliqués dans la résistance aux insecticides chez les insectes autres que *Drosophila melanogaster* et *Musca domestica* ou bien chez les arachnides:

- La mise en évidence, chez *C. pipiens*, de deux formes d'AChE possédant des activités catalytiques distinctes supporte les deux hypothèses et l'analyse biochimique de ces AChE n'a pas permis de déterminer la nature de l'AChE impliquée dans la résistance aux insecticides (Bourguet et al., J. Neurochemistry, 1996, 67, 2115-2123).

- Un deuxième gène *ace* a été isolé chez les arachnides, toutefois ce gène n'est pas impliqué dans la résistance aux insecticides (Hernandez et al., Baxter et al., précités).

- Un deuxième gène *ace* n'a pu être isolé chez les insectes malgré de nombreuses tentatives dans différentes espèces (Menozzi et al., Tomita et al., Mori et al., précités ; Severson et al., J. Hered., 1997, 88, 520-524).

Il ressort de ce qui précède que la nature du gène et de la cible (AChE) impliqués dans la résistance aux organophosphorés et aux carbamates n'ont pas été identifiés chez la plupart des insectes et chez les arachnides, notamment chez ceux où ils ont été recherchés ; on peut citer les plus importants dans les domaines de la santé humaine ou animale et de l'agriculture comme les vecteurs de pathogènes et les nuisibles, notamment de nombreux moustiques comme *Culex pipiens*, *Aedes aegypti*, *Anopheles gambiae*, *Anopheles albimanus*, *Anopheles stephensi*, et des ravageurs des cultures comme *Aphis gossypii*, *Nephotettix cincticeps* et *Leptinotarsa decemlineata*.

Les Inventeurs ont identifié un nouveau locus du gène *ace* dans le génome d'*Anopheles gambiae* et de 15 espèces différentes de moustiques et ils ont montré que ce nouveau locus, non-homologue au locus précédemment décrit chez *D. melanogaster*, était impliqué dans la résistance aux insecticides chez les moustiques.

Ce nouveau gène représente un outil de diagnostic pour la détection génétique de la résistance aux insecticides (organophosphorés, carbamates) dans les populations de moustiques. L'AChE codée par ce gène représente une cible pour le criblage de nouvelles molécules actives sur les populations de moustiques résistants aux insecticides actuellement utilisés.

La présente invention a, en conséquence, pour objet une protéine, caractérisée en ce qu'elle comprend une région catalytique centrale qui présente une séquence en acides aminés sélectionnée dans le groupe constitué par la séquence SEQ

ID NO: 1 et les séquences présentant au moins 60 % d'identité ou 70 % de similarité avec la séquence SEQ ID NO: 1, à l'exclusion de la séquence NCBI AAK0973 correspondant à l'acétylcholinestérase de *Schizaphis graminum*.

La protéine selon l'invention représente une nouvelle acétylcholinestérase d'insecte, dénommée ci-après AchE1, responsable de la résistance aux organophosphorés et aux carbamates, au moins chez les moustiques, notamment chez *C. pipiens* ; le locus codant pour ladite AchE1 est dénommée ci-après *ace-1* ; *ace-2* représente le second locus *ace*, qui n'est pas impliqué dans la résistance aux insecticides chez les moustiques. L'unique gène *ace* présent dans *Drosophila melanogaster*, qui est homologue à *ace-2*, est donc également dénommé *ace-2*.

Conformément à l'invention, ladite région catalytique centrale contient le domaine catalytique de l'AChE et correspond à celle située entre les positions 70 et 593 de la séquence de l'AChE1 d'*Anopheles gambiae* (SEQ ID NO: 3, 643 acides aminés) ; elle correspond à celle située respectivement entre les positions 100 et 629 de la séquence d'AChE1 de *Schizaphis graminum* (NCBI AAK0973), 60 et 582 de la séquence de l'AChE1 de *Culex pipiens* (SEQ ID NO: 7), 34 et 593 de la séquence d'AChE2 d'*Anopheles gambiae* (figure 1, SEQ ID NO: 53), et 41 et 601 de la séquence d'AChE2 de *Drosophila melanogaster* (NCBI AAF54915). Cette région centrale qui contient le domaine catalytique est conservée chez les vertébrés et les invertébrés alors que les extrémités N- et C-terminales présentent une forte variabilité entre les différentes espèces.

Conformément à l'invention, l'identité d'une séquence par rapport à une séquence de référence (SEQ ID NO: 1) s'apprécie en fonction du pourcentage de résidus d'acides aminés qui sont identiques, lorsque les séquences correspondant à la région catalytique telle que définie ci-dessus sont alignées, de manière à obtenir le maximum de correspondance entre elles.

Une protéine ayant une séquence en acides aminés ayant au moins X % d'identité avec la séquence de référence SEQ ID NO: 1 est définie, dans la présente invention comme une protéine dont la séquence correspondant à la région catalytique centrale telle que définie ci-dessus peut inclure jusqu'à 100-X altérations pour 100 acides aminés de la séquence SEQ ID NO: 1. Au sens de la présente invention, le terme altération inclut les délétions, les substitutions ou les insertions consécu-



tives ou dispersées d'acides aminés dans la séquence de référence. Cette définition s'applique, par analogie, aux molécules d'acide nucléique.

La similarité d'une séquence par rapport à la séquence de référence SEQ ID NO 1 s'apprécie en fonction du pourcentage de résidus d'acides aminés qui sont identiques ou qui diffèrent par des substitutions conservatives, lorsque les séquences correspondant à la région catalytique centrale telle que définie ci-dessus sont alignées de manière à obtenir le maximum de correspondance entre elles. Au sens de la présente invention, on entend par substitution conservative, la substitution d'un acide aminé par un autre qui présente des propriétés chimiques similaires (taille, charge ou polarité), qui généralement ne modifie pas les propriétés fonctionnelles de la protéine.

Une protéine ayant une séquence en acides aminés ayant au moins X % de similarité avec la séquence SEQ ID NO: 1 est définie, dans la présente invention comme une protéine dont la séquence correspondant à la région catalytique centrale telle que définie ci-dessus peut inclure jusqu'à 100-X altérations non-conservatives pour 100 acides aminés de la séquence de référence. Au sens de la présente invention, le terme altérations non-conservatives inclut les délétions, les substitutions non-conservatives ou les insertions consécutives ou dispersées d'acides aminés dans la séquence SEQ ID NO: 1.

La comparaison de l'ACHé1 selon l'invention avec les AChE d'insecte disponibles sur les bases de données, par alignement des séquences correspondant à la région centrale telle que définie ci-dessus, à l'aide du logiciel BLAST (<http://www.ncbi.nlm.nih.gov/gorf/bl2.html>, paramètres par défaut, filtre inactivé) montre que :

- les séquences d'ACHé1 et d'ACHé2 d'insecte présentent 36-39% d'identité (53-57% similarité) entre elles.
- les séquences d'ACHé1 d'insecte présentent 65-97% d'identité (79-98% similarité) entre elles,
- les séquences d'ACHé2 d'insecte présentent 58-99% d'identité (73-99% similarité) entre elles,

En outre, l'analyse phylogénétique des AChE des différentes espèces animales montre que les séquences protéiques d'ACHé1 forment un groupe autonome

significatif (bootstrap 795/1000), et que les AChE1 d'insecte forment un sous-groupe distinct significatif (bootstrap 856/1000).

L'AChE1 selon l'invention comprend des motifs caractéristiques des AChE (figure 1) situés aux positions suivantes, respectivement dans la séquence SEQ ID NO: 3 et dans la séquence de référence de *Torpedo marmorata* (SWISSPROT P07962) : un motif canonique du type FGESAG autour de la sérine en position 266 (200), qui est caractéristique du site actif des AChE, un site de liaison à la choline (résidu Tryptophane en position 151 (84)), trois résidus de la triade catalytique (résidus sérine, acide glutamique et histidine, respectivement en positions 266 (200), 392 (327) et 506 (440)), six résidus cystéine potentiellement impliqués dans des ponts disulfures conservés (C<sub>134(67)</sub>-C<sub>161(94)</sub>; C<sub>320(254)</sub>-C<sub>333(265)</sub>; C<sub>468(402)</sub>-C<sub>589(521)</sub>), des résidus aromatiques bordant la gorge du site actif (10 résidus) et un résidu phénylalanine en position 355 (290) mais pas en position 353 (288), qui distingue les AChE d'invertébrés de celles de vertébrés. Elle possède également un peptide C-terminal hydrophobe correspondant à un signal d'addition d'un glycolipide, indiquant le clivage post-traductionnel d'un fragment C-terminal et l'addition d'un résidu d'ancrage glycolipidique comme chez *Drosophila* ; le résidu cystéine dans la séquence C-terminale précédant le site potentiel de clivage du peptide hydrophobe pourrait être impliqué dans une liaison disulfure intermoléculaire, liant les deux sous-unités catalytiques du dimère d'AChE.

L'AChE1 selon l'invention se distingue de l'AChE de *Drosophila* (AChE2) par l'absence d'une insertion hydrophile de 31 acides aminés entre les résidus situés aux positions 174 et 175 de la séquence SEQ ID NO: 3 (figure 1) ; cette insertion hydrophile pourrait être caractéristique de l'AChE2, au moins chez les diptères.

L'invention englobe les AChE1 d'insecte sensibles ou résistantes aux organophosphorés et aux carbamates.

Au sens de la présente invention on entend par "AChE sensible", une AChE dont l'activité acétylcholinestérase est inhibée en présence d'organophosphorés ou de carbamates.

Au sens de la présente invention on entend par "AChE résistante", une AChE dont l'activité n'est pas inhibée par des concentrations en organophosphorés ou en carbamates qui inhibent 100 % de l'activité de "l'AChE sensible" correspondante

issue d'un individu de la même espèce ; cette "AChE résistante" diffère de la précédente par la présence d'une ou plusieurs mutations dans sa séquence en acides aminés (substitutions d'acides aminés) qui modifient sa sensibilité aux inhibiteurs de l'acétylcholinestérase ; parmi ces mutations on peut citer les suivantes F78S, I129V, G227A, F288Y, les acides aminés étant numérotés en référence à la séquence de l'AChE de *Torpedo marmorata* (SWISSPROT P07962).

L'activité acétylcholinestérase et les paramètres catalytiques des AChE sont mesurés par les techniques enzymatique classiques telles que celles décrites dans Bourguet et al., précité.

Les protéines selon l'invention incluent toute protéine naturelle, synthétique, semi-synthétique ou recombinante de n'importe quel organisme procaryote ou eucaryote, comprenant ou consistant en une séquence d'acides aminés d'une protéine AChE1 telle que définie ci-dessus. Elles incluent notamment les protéines naturelles isolées chez n'importe quelle espèce d'insecte, ainsi que les protéines recombinantes produites dans un système d'expression approprié.

Selon un mode de réalisation avantageux de ladite AChE1, elle correspond à celle d'un insecte qui appartient à l'ordre des diptères (*Diptera*) ; de manière préférée, ledit insecte est choisi dans la famille des *Culicidae*, parmi les genres *Culex*, *Aedes* et *Anopheles*.

Selon une disposition avantageuse de ce mode de réalisation, ladite AChE1 est constituée par les séquences SEQ ID NO: 3, SEQ ID NO: 5 et SEQ ID NO: 7, correspondant à la séquence complète de l'AChE1 de respectivement deux souches d'*Anopheles gambiae* et une souche de *Culex pipiens*, sensibles aux organophosphorés et aux carbamates.

Selon une autre disposition avantageuse de ce mode de réalisation, ladite région centrale catalytique de l'AChE1 comprend une séquence sélectionnée dans le groupe constitué par les séquences SEQ ID NO: 8 à 21 représentant un fragment d'environ 91 acides aminés (fragment K, figure 1), correspondant à celui situé entre les positions 445 et 535 de la séquence SEQ ID NO: 3.

La présente invention a également pour objet un peptide, caractérisé en ce qu'il est constitué par un fragment d'au moins 7 acides aminés de la protéine

AChE1, telle que définie ci-dessus ; ces fragments sont particulièrement utiles pour la production d'anticorps reconnaissant spécifiquement la protéine AChE1.

La présente invention a également pour objet des anticorps, caracté-  
risés en ce qu'ils sont dirigés contre la protéine AChE1 ou un fragment de celle-ci, tels  
5 que définis ci-dessus.

---

Conformément à l'invention, lesdits anticorps sont soit des anticorps  
monoclonaux, soit des anticorps polyclonaux.

Ces anticorps peuvent être obtenus par les méthodes classiques,  
connues en elles-mêmes, comprenant notamment l'immunisation d'un animal avec  
10 une protéine ou un peptide conforme à l'invention, afin de lui faire produire des  
anticorps dirigés contre ladite protéine ou ledit peptide.

La présente invention a également pour objet une molécule d'acide  
nucléique isolée, caractérisée en ce qu'elle présente une séquence sélectionnée dans le  
groupe constitué par :

- 15 - les séquences codant pour une protéine AChE1 telle que définie ci-  
dessus (ADNc et fragment d'ADN génomique correspondants au gène *ace-1*), et
- les séquences complémentaires des précédentes, sens ou anti-sens.
- les fragments d'au moins 8 pb, de préférence de 15 pb à 500 pb des  
séquences précédentes.

20 L'invention englobe, les séquences des allèles du gène *ace-1* issues  
de n'importe quel insecte, ainsi que les séquences des mutants naturels (allèles sensi-  
bles et résistants) ou artificiels du gène *ace-1* codant pour une protéine AChE1 sensi-  
ble ou résistante, telle que définie ci-dessus.

Selon un mode de réalisation avantageux de l'invention, ladite  
25 séquence codant pour une protéine AChE1 est sélectionnée dans le groupe constitué  
par :

- a) les séquences SEQ ID NO: 2, SEQ ID NO: 4 et SEQ ID NO: 6 qui  
correspondent à l'ADNc de la protéine AChE1 de séquence en acides aminés, respecti-  
vement SEQ ID NO: 3, SEQ ID NO: 5 et SEQ ID NO: 7, telles que définie ci-dessus,  
30 et

b) les séquences SEQ ID NO: 22 et SEQ ID NO: 23 qui correspondent au gène *ace-1* d'*Anopheles gambiae* codant les AChE1 en a), lequel gène présente une organisation exon-intron comprenant au moins 7 exons (Tableau I).

**Tableau I : Organisation Intron-Exon du gène *ace-1***

	Site 5'		Site 3'	
	Position	Séquence	Position	Séquence
Intron 1	nd	nd	1179	ttcag/ACGCA
Intron 2	1315	CTCGG/gtaag	1400	ggcag/ACGCG
Intron 3	1938	CTACG/gtagg	2017	gtcag/CTGGG
Intron 4	2214	CTAAG/gtacg	2301	tccag/AGCAC
Intron 5	3009	ACCGG/gtaag	3075	tacag/CAATC
Intron 6	3248	TACCT/gtaag	3355	aacag/CGAAC

5

Selon un autre mode de réalisation avantageux de l'invention, ledit fragment est sélectionné dans le groupe constitué par les amorces de séquence SEQ ID NO: 39 à 50 et les fragments de séquences SEQ ID NO: 24 à 38.

Les molécules d'acide nucléique selon l'invention sont obtenues par les méthodes classiques, connues en elles-mêmes, en suivant les protocoles standards tels que ceux décrits dans *Current Protocols in Molecular Biology* (Frederick M. AUSUBEL, 2000, Wiley and son Inc, Library of Congress, USA). Par exemple, elles peuvent être obtenues par amplification d'une séquence nucléique par PCR ou RT-PCR, par criblage de banques d'ADN génomique par hybridation avec une sonde  
15 homologue, ou bien par synthèse chimique totale ou partielle.

Les molécules d'acides nucléiques telles que définies ci-dessus peuvent être utilisées comme sondes ou comme amorces pour isoler le gène *ace-1* d'autres espèces ou des allèles de ce gène, notamment par criblage d'une banque d'ADN génomique ou d'ADNc, ainsi que pour détecter/amplifier des molécules d'acide nucléique  
20 (ARNm ou ADN génomique) codant une protéine AChE1 telle que définie ci-dessus.

Ces différentes molécules d'acides nucléiques permettent de mettre en évidence le gène *ace-1*, des variants alléliques de ce gène, une altération fonctionnelle de ce gène *ace-1* (changement substantiel de la sensibilité aux insecticides) résultant d'une mutation (insertion, délétion ou substitution) d'un ou plusieurs nucléotides au niveau dudit gène.  
25

La présente invention a également pour objet une méthode de détection d'insectes porteurs d'une résistance aux insecticides de la classe des organophosphorés et des carbamates, caractérisée en ce qu'elle comprend :

- la préparation d'un échantillon d'acides nucléiques à partir d'insectes à tester, et

---

~~la détection par tout moyen approprié de la présence, dans ledit échantillon d'acides nucléiques, d'une mutation dans le gène *ace-1* tel que défini ci-dessus.~~

Ladite détection est réalisée par les techniques classiques qui sont connues en elles mêmes, par exemple : (i) par amplification d'une région dudit gène *ace-1* susceptible de contenir une mutation, puis détection de ladite mutation par séquençage ou par digestion par une enzyme de restriction appropriée, du produit de PCR obtenu, ou bien (ii) par hybridation avec une sonde marquée spécifique d'une région dudit gène *ace-1* susceptible de contenir une mutation, puis détection directe des mésappariements et/ou digestion par une enzyme de restriction appropriée.

De préférence un fragment d'environ 320 pb (fragment K) est amplifié à l'aide des amorces SEQ ID NO: 39 et SEQ ID NO: 40. Par exemple, chez les moustiques on obtient un fragment de séquence SEQ ID NO: 24 à 38 qui présente des mutations entre les moustiques sensibles et résistants aux insecticides. Par exemple, chez *C. pipiens* on observe 3 substitutions dans la séquence des individus résistants dont l'une introduit un site *EcoRI*. L'analyse du profil de restriction après amplification PCR du fragment K et digestion des produits obtenus par *EcoRI* (analyse RFLP), permet de détecter rapidement le génotype *ace-1* dans une population de *C. pipiens* ; la présence d'un seul fragment correspond aux homozygotes résistants (RR), la présence de 2 fragments d'environ 106 pb et 214 pb correspond aux individus homozygotes sensibles (SS) et la présence de 3 fragments de 106 pb, 214 pb et 320 pb correspond aux individus hétérozygotes résistants (RS).

La présente invention a également pour objet un réactif de détection d'insectes porteurs d'une résistance aux organophosphorés et aux carbamates, caractérisé en ce qu'il est sélectionné dans le groupe constitué par : les molécules d'acide nucléique et leurs fragments tels que définis ci-dessus (sondes , amorces) et les anticorps tels que définis ci-dessus.

La présente invention a également pour objet un vecteur recombinant, caractérisé en ce qu'il comprend un insert sélectionné dans le groupe constitué par les molécules d'acides nucléiques codant une protéine AChE1 et leurs fragments tels que définis ci-dessus.

5 De préférence, ledit vecteur recombinant est un vecteur d'expression dans lequel ladite molécule d'acide nucléique ou l'un de ses fragments sont placés sous le contrôle d'éléments régulateurs de la transcription et de la traduction appropriés.

Ces vecteurs sont construits et introduits dans des cellules hôtes par les méthodes classiques d'ADN recombinant et de génie génétique, qui sont connues  
10 en elles-mêmes. De nombreux vecteurs dans lesquels on peut insérer une molécule d'acide nucléique d'intérêt afin de l'introduire et de la maintenir dans une cellule hôte eucaryote ou procaryote, sont connus en eux-mêmes ; le choix d'un vecteur approprié dépend de l'utilisation envisagée pour ce vecteur (par exemple répllication de la  
15 séquence d'intérêt, expression de cette séquence, maintien de la séquence sous forme extrachromosomique ou bien intégration dans le matériel chromosomique de l'hôte), ainsi que de la nature de la cellule hôte. Par exemple, on peut utiliser des vecteurs viraux comme les baculovirus ou non-viraux comme des plasmides. Pour exprimer l'AChE1, l'ADNc d'*ace-1* peut être placé sous le contrôle d'un promoteur constitutif  
20 comme le promoteur de l'actine 5C, dans un vecteur approprié et ledit vecteur recombinant est introduit dans des cellules d'insecte telles que des cellules de drosophile (cellules de Schneider S2 ).

La présente invention a également pour objet des cellules procaryotes ou eucaryotes, modifiées par un vecteur recombinant tel que défini ci-dessus ; de préférence ces cellules sont des cellules d'insectes.

25 Les vecteurs recombinants et les cellules modifiées telles que définies ci-dessus, sont utiles notamment pour la production des protéines et des peptides AChE1 selon l'invention.

La présente invention a également pour objet un animal invertébré transgénique, caractérisé en ce qu'il contient des cellules modifiées par au moins une  
30 molécule d'acide nucléique telle que définie ci-dessus ; de préférence ledit animal est un insecte.

Les animaux transgéniques et les cellules modifiées telles que définis ci-dessus, sont utiles notamment pour le criblage de substances insecticides.

La présente invention a également pour objet une méthode de criblage d'une substance insecticide, caractérisée en ce qu'elle comprend :

5 a) la mise en contact de la substance à tester avec une protéine

~~ACHEl sélectionnée parmi la protéine AChE1 isolée selon l'invention, un extrait de~~  
cellules modifiées ou un échantillon biologique d'un animal transgénique contenant ladite protéine AChE1, tels que définis ci-dessus, en présence d'acétylcholine ou de l'un de ses dérivés,

10 b) la mesure par tout moyen approprié, de l'activité acétylcholinestérase du mélange obtenu en a), et

c) la sélection des substances capables d'inhiber ladite activité.

La présente invention a également pour objet une méthode de criblage d'une substance insecticide, caractérisée en ce qu'elle comprend :

15 - la mise en contact d'un animal transgénique tel que défini ci-dessus, avec la substance à tester, et

- la mesure de la survie de l'animal.

Avantageusement, lesdites méthodes de criblages mettent en œuvre des AChE1 résistantes aux organophosphorés ou aux carbamates ou bien des cellules  
20 ou des animaux transgéniques les contenant.

La présente invention a également pour objet un réactif de criblage de substances insecticides, caractérisé en ce qu'il est sélectionné dans le groupe constitué par les protéines AChE1, les vecteurs recombinants, les cellules modifiées et les animaux transgéniques tels que définis ci-dessus.

25 Des substances insecticides capables d'inhiber l'activité acétylcholinestérase des protéines AChE1 résistantes aux insecticides de la classe des organophosphorés et des carbamates couramment utilisés ont des applications : en santé humaine et animale, pour lutter contre les vecteurs de pathogènes (par exemple *Aedes aegypti*, vecteur d'arboviroses comme la dengue et la fièvre jaune, *Culex pipiens* vecteur du virus West-Nile, *Anopheles gambiae* vecteur africain de l'agent du paludisme, etc) et dans le  
30 domaine de l'agriculture, pour lutter contre les insectes nuisibles qui dévastent les récoltes (par exemple le doryphore (*Leptinotarsa decemlineata*) qui s'attaque aux



pommes-de-terre, les pucerons ravageurs comme *Aphis gossypii* et *Myzus persicae*, etc.).

L'invention a en outre pour objet une trousse de détection et/ou de criblage pour la mise en œuvre des méthodes telles que définies ci-dessus, caractérisée en ce qu'elle inclut au moins un réactif tel que défini ci-dessus.

Outre les dispositions qui précèdent, l'invention comprend encore d'autres dispositions qui ressortiront de la description qui va suivre, qui se réfère à des exemples de mise en œuvre du gène *ace-1* et de ses produits (ADNc, protéine) selon la présente invention ainsi qu'au tableau résumant les séquences de la Demande et aux  
10 dessins annexés dans lesquels :

- la figure 1 illustre l'alignement des séquences en acides aminés des protéines AChE1 d'*Anopheles gambiae*, *Schizaphis graminum*, *An. stephensi*, *Aedes aegypti*, *Drosophila melanogaster*, *Lucilia cuprina*, *Musca domestica* et *Culex pipiens*. Par convention, les acides aminés sont numérotés en référence à la séquence  
15 de l'AChE du poisson torpille (*Torpedo marmorata*; SWISSPROT P07962). Les séquences N- et C- terminales ne sont pas représentées en raison de leur variabilité. Les acides aminés conservés entre AChE1 et AChE2 sont indiqués en gris. Les acides aminés spécifiques d'AChE2 sont indiqués en noir. Les 3 résidus représentant la triade catalytique (S<sub>200</sub>, E<sub>327</sub> et H<sub>440</sub>) sont encadrés. Le site de liaison à la choline (W<sub>84</sub>) est  
20 souligné. Les cercles représentent la position des 14 résidus aromatiques bordant la gorge du site actif dans l'AChE de *Torpedo*, dont 10 sont présents dans toutes les AChE1 et AChE2 (cercles pleins), les autres n'étant pas conservés (cercles vides). Trois liaisons disulfures intramoléculaires entre des résidus cystéines sont indiquées. La flèche horizontale indique la position du fragment K (amplifié à l'aide des amorces  
25 PdirAGSG et PrevAGSG). La région hypervariable d'AChE2 qui est absente dans AChE1 est entourée.

- la figure 2 illustre la détection génétique des moustiques résistants aux organophosphorés et aux carbamates par PCR-RFLP :

- la figure 2 A représente la comparaison de la séquence en acides aminés du fragment K de différentes espèces de moustiques: Cx Pip (*Culex pipiens*),  
30 Ae alb (*Aedes albopictus*), Ae aeg (*Aedes aegypti*), An alb (*Anopheles albimanus*), An gamb (*Anopheles gambiae*), An fun (*Anopheles funestus*), An nil (*Anopheles nili*), An

sac (*Anopheles sacharovi*), An pse (*Anopheles pseudopunctipennis*). Les acides aminés variants sont grisés. Les séquences suivantes sont identiques: *An. darlingi* et *An. albimanus*; *An. sundaicus*, *An. gambiae* et *An. arbiensis*; *An. moucheti*, *An. funestus* et *An. minimus*; *An. stephensi* et *An. saccharovi*.

5 . la figure 2B illustre la comparaison des séquences nucléotidiques

correspondant au fragment K des souches sensibles (S-LAB) et résistantes (SR). Les nucléotides variants sont grisés (t → c en position 3 ; a → g en position 84 : le site *EcoRI* (gaattc) situé autour de cette position, utilisé pour l'analyse PCR-RFLP, est présent uniquement dans la souche S-LAB ; c → t en position 173). La figure 2C illustre les profils de restriction obtenus après électrophorèse en gel d'agarose des produits de digestion par *EcoRI*, du fragment K amplifié par PCR. La souche homozygote sensible S-LAB présente un profil caractérisé par 2 bandes (214 pb et 106 pb), la souche homozygote résistante présente un profil caractérisé par une seule bande de 320 pb et les moustiques résistants issus du croisement en retour présentent un profil hétérozygote caractérisé par 3 bandes (320 pb, 214 pb et 106 pb).

15 - la figure 3 illustre l'arbre phylogénétique des protéines AChE. L'analyse phylogénétique a été réalisée à partir de 47 séquences de protéines AChE de 35 espèces différentes provenant de la base de données ESTHER (<http://www.ensam.inra.fr/cgi-bin/ace/index>). Les séquences ont été alignées et un arbre a été construit comme décrit à l'exemple 1. Seuls les nœuds correspondant à des valeurs de "bootstrap" > 50% (c'est à dire des scores supérieurs à 500) sont indiqués. L'échelle représente une divergence de 10 %. Agam: *An. gambiae* ; Aeg: *Aedes aegypti* ; Aste: *Anopheles stephensi*; Cpip: *Culex pipiens*; Dmel: *Drosophila melanogaster*; Lcup: *Lucilia cuprina*; Mdom: *Musca domestica*; Ldec: *Leptinotarsa decemlineata*; Ame1: *Apis mellifera*; Ncin: *Nephotettix cincticeps* ; Sgra: *Schizaphis graminum* ; Rapp: *Rhipicephalus appendiculatus* ; Bmic: *Boophilus microplus*; Bdec: *Boophilus decoloratus*; Hsap: *Homo sapiens* ; Btau: *Bos taurus* ; Fcat: *Felis catus* ; Ocun: *Oryctolagus cuniculus* ; Rnor : *Rattus norvegicus* ; Mmus: *Mus musculus* ; Ggal: *Gallus gallus* ; Drer: *Danio rerio* ; Eele: *Electrophorus electricus* ; Tamr: *Torpedo marmorata* ; Tcal: *Torpedo californica* ; Bfas: *Bungarus fasciatus* ; Mglu: *Myxine glutinosa* ; Bflo: *Branchiostoma floridae* ; Blan: *Branchiostoma lanceolatum* ;

Cint: *Ciona intestinalis* ; Csav: *Ciona savignyi* ; Cele: *Caenorhabditis elegans* ; Cbrig: *Caenorhabditis briggsae* ; Dviv: *Dictyocaulus viviparus*; Lopa: *Loligo opalescens*.

- la figure 4 illustre le cladogramme des protéines AChE1 et AChE2.

Les séquences des protéines AChE1 et AChE2 ont été traitées comme à la figure 1. La

5 séquence Bmic a été ajoutée comme séquence externe pour définir l'origine de l'arbre.

Les cadres marqués d'une astérisque représentent les protéines codées par un gène qui ségrège avec la résistance aux insecticides. Les cadres vides représentent les protéines codées par un gène qui ne ségrège pas avec la résistance aux insecticides. L'échelle correspond à une divergence de 10 %.

10

**Tableau II: Liste des séquences**

Numéro d'identification	Séquence
SEQ ID NO: 1	fragment de la région centrale de la protéine AChE1 <i>Anopheles gambiae</i> (positions 70 à 593 de la SEQ ID NO: 3).
15 SEQ ID NO:2	ADNc AChE1 <i>Anopheles gambiae</i>
SEQ ID NO: 3	Protéine AChE1 <i>Anopheles gambiae</i>
SEQ ID NO: 4	ADNc AChE1 <i>Anopheles gambiae</i> (souche KISUMU)
SEQ ID NO: 5	Protéine AChE1 <i>Anopheles gambiae</i> (souche KISUMU)
SEQ ID NO: 6	ADNc AChE1 <i>Culex pipiens</i> (souche S-LAB)
20 SEQ ID NO: 7	Protéine AChE1 <i>Culex pipiens</i> (souche S-LAB)
SEQ ID NO: 8	fragment peptidique K AChE1 <i>Culex pipiens</i>
SEQ ID NO: 9	fragment peptidique K AChE1 <i>Aedes aegypti</i>
SEQ ID NO: 10	fragment peptidique K AChE1 <i>Aedes albopictus</i>
SEQ ID NO: 11	fragment peptidique K peptidique AChE1 <i>Anopheles darlingi</i>
25 SEQ ID NO: 12	fragment peptidique K AChE1 <i>An. sundaicus</i>
SEQ ID NO: 13	fragment peptidique K AChE1 <i>An. minimus</i>
SEQ ID NO: 14	fragment peptidique K AChE1 <i>An. moucheti</i>
SEQ ID NO: 15	fragment peptidique K AChE1 <i>An. arabiensis</i>
SEQ ID NO: 16	fragment peptidique K AChE1 <i>An. funestus</i>
30 SEQ ID NO: 17	fragment peptidique K AChE1 <i>An. pseudopunctipennis</i>
SEQ ID NO: 18	fragment peptidique K AChE1 <i>An. sacharovi</i>
SEQ ID NO: 19	fragment peptidique K AChE1 <i>An. stephensi</i>
SEQ ID NO: 20	fragment peptidique K AChE1 <i>An. albimanus</i>
SEQ ID NO: 21	fragment peptidique K AChE1 <i>An. nili</i>
35 SEQ ID NO: 22	gène <i>ace-1</i> <i>An. gambiae</i>
SEQ ID NO: 23	gène <i>ace-1</i> <i>An. gambiae</i> KISUMU
SEQ ID NO: 24	fragment nucléotidique K AChE1 <i>C. pipiens</i> (souche S-LAB)
SEQ ID NO: 25	fragment nucléotidique K AChE1 <i>C. pipiens</i> (souche SR)
SEQ ID NO: 26	fragment nucléotidique K AChE1 <i>Aedes aegypti</i>
40 SEQ ID NO: 27	fragment nucléotidique K AChE1 <i>Aedes albopictus</i>
SEQ ID NO: 28	fragment nucléotidique K AChE1 <i>Anopheles darlingi</i>
SEQ ID NO: 29	fragment nucléotidique K AChE1 <i>An. sundaicus</i>
SEQ ID NO: 30	fragment nucléotidique K AChE1 <i>An. minimus</i>

5	SEQ ID NO: 31	fragment nucléotidique K AChE1 <i>An. moucheti</i>
	SEQ ID NO: 32	fragment nucléotidique K AChE1 <i>An. arabiensis</i>
	SEQ ID NO: 33	fragment nucléotidique K AChE1 <i>An. funestus</i>
	SEQ ID NO: 34	fragment nucléotidique K AChE1 <i>An. pseudopunctipennis</i>
	SEQ ID NO: 35	fragment nucléotidique K AChE1 <i>An. sacharovi</i>
	SEQ ID NO: 36	fragment nucléotidique K AChE1 <i>An. stephensi</i>
	SEQ ID NO: 37	fragment nucléotidique K AChE1 <i>An. albimanus</i>
	SEQ ID NO: 38	fragment nucléotidique K AChE1 <i>An. nili</i>
10	SEQ ID NO: 39	amorçe PkdirAGSG
	SEQ ID NO: 40	amorçe PkrevAGSG
	SEQ ID NO: 41	amorçe PkdirAGSG
	SEQ ID NO: 42	amorçe PbrevAGSG
15	SEQ ID NO: 43	amorçe culex-bdir1
	SEQ ID NO: 44	amorçe culex-krev1
	SEQ ID NO: 45	amorçe AG1-Adir
	SEQ ID NO: 46	amorçe AG1-Arev
20	SEQ ID NO: 47	amorçe AG1-Bdir
	SEQ ID NO: 48	amorçe AG1-Brev
	SEQ ID NO: 49	amorçe AG1-Cdir
	SEQ ID NO: 50	amorçe AG1-Crev
	SEQ ID NO: 51	Protéine AChE1 <i>Ciona intestinalis</i>
	SEQ ID NO: 52	Protéine AChE1 <i>Ciona savignyi</i>
	SEQ ID NO: 53	Protéine AChE2 <i>Anopheles gambiae</i>

## 25 **EXEMPLE 1 : Matériels et méthodes**

### a) Souches et croisements

Cinq souches de *C. pipiens* ont été utilisées: S-LAB, une souche standard sensible aux insecticides (Georghiou et al., 1966, Bull. Wld. Hlth Org., 35, 691-708), SA1, SA4 et EDIT qui possèdent une seule AChE sensible aux insecticides, et SR qui est homozygote pour une AChE insensible aux insecticides (Berticat et al., 30 Genet. Res., 2002, 79, 41-47). Les souches possédant une AChE sensible et insensible sont dénommées respectivement S et R.

### b) Nomenclature des gènes *ace* et numérotation des séquences d'acides aminés

*ace-1* représente le locus codant pour une AChE cholinergique responsable de la résistance aux organophosphorés et aux carbamates chez *C. pipiens* (AChE1), précédemment dénommé *Ace.1* (Raymond et al., Genetica, 2001, 112/113, 287-296). *ace-2* représente le second locus *ace*, qui n'est pas impliqué dans la résistance aux insecticides chez *C. pipiens* (précédemment dénommé *Ace.2*), dont la fonction est inconnue chez *C. pipiens*. L'unique gène *ace* présent dans *Drosophila* 40 *melanogaster*, qui est homologue à *ace-2*, est donc dénommé de même.

Dans les analyses qui suivent, les positions des résidus d'acides aminés sont indiquées en référence à la séquence de l'AChE du poisson torpille [*Torpedo marmorata*; GENBANK P07962], selon la nomenclature recommandée par Massoulié et al., 1992, *In Multidisciplinary approaches to cholinesterase functions*, eds, Schafferman, A. & Velan, B. (Plenum Press New York), p 285-288].

c) Analyse de la transmission du gène *ace-1*

Les femelles étant indiquées en premier, des croisements F1 (S X R) et des croisements en retour (F1 X S-LAB et S-LAB X F1) ont été obtenus par croisement en masse d'adultes. Quelques larves issues des croisements en retour ont été traitées avec une dose de carbamate (propoxur, 4mg/L) qui tue 100 % des larves sensibles. La liaison entre *ace-1* et la résistance au propoxur a été étudiée par RFLP chez les larves survivantes, à partir d'un produit PCR de 320 pb permettant d'identifier les allèles S et R. Les expériences ont été réalisées de façon indépendante, avec S = SA1, S = SA4 et S = EDIT.

d) Analyse des séquences et assemblage des gènes

Toutes les analyses de séquences ont été effectuées à partir des séquences brutes d'*Anopheles gambiae* disponibles sur le serveur INFOBIOGEN (<http://www.infobiogen.fr>) et des outils disponibles sur le site (<http://www.ncbi.nlm.nih.gov/blast/blast>). Les séquences génomiques codant une AChE ont été identifiées à l'aide des logiciels TBLASTN et BLAST (Altschul et al., J. Biol. Mol., 1990, 215, 403-410). Les séquences génomiques identifiées ont été assemblées à l'aide du logiciel ABI Prism Auto-Assembler (v2.1, PERKIN ELMER). Les séquences ont été vérifiées et corrigées à l'aide du logiciel Ensembl Trace Server (<http://trace.ensembl.org/>). Deux concaténations de respectivement 5195 et 6975 paires de bases, codant respectivement pour AChE1 et AChE2 ont été assemblées à partir de respectivement 64 et 74 séquences indépendantes (redondance moyenne de 10,5 et 6,5). Les exons et les séquences protéiques ont été identifiés à l'aide d'une combinaison entre les logiciels FGENESH (<http://www.sanger.uk>) et BLASTX (<http://www.ncbi.nlm.nih.gov>). Les séquences génomiques d'AChE d'ascidies ont été assemblées à partir de séquences brutes déposées dans les bases de données du NCBI (*Ciona savignyi*) et du Doe Joint Institute (*Ciona intestinalis*, [http://www.jgi.doe.gov/programs/ciona/ciona\\_mainage.html](http://www.jgi.doe.gov/programs/ciona/ciona_mainage.html)). Les recherches dans les

bases de données de *Drosophila* ont été effectuées à l'aide de Flybase (<http://www.fruitfly.org/>).

#### e) Comparaisons de séquences

Les séquences des protéines AChE1 et AChE2 d'*Anopheles gambiae* déduites des séquences génomiques et les séquences peptidiques déduites de fragments PCR de *C. pipiens* et *A. aegypti* ont été alignées avec celles des AChE connues, à l'aide du logiciel ClustalW, en utilisant une matrice BLOSUM et des paramètres par défaut (Thompson et al., N.A.R., 1994, 22, 4673-4680).

Un arbre phylogénétique a été construit en utilisant l'algorithme du plus proche voisin (*neighbour-joining algorithm*) de la version DDBJ de Clustal W ([http://hypernig.nig.ac.jp/homology/ex\\_clustalw-e.shtml](http://hypernig.nig.ac.jp/homology/ex_clustalw-e.shtml)). L'analyse de type *Bootstrap* (1000 comptages et 111 valeurs d'entrée) a été utilisée pour évaluer les degrés de confiance pour la topologie de l'arbre. La construction des arbres a été réalisée à l'aide du logiciel Treeview (v1.6.6).

#### f) Numéros d'accension

Les numéros des séquences (numéros d'accension dans les bases de données ou les numéros d'identification dans la liste de séquences) ayant servi à l'analyse génétique sont les suivants.

- Craniata : *Homo sapiens* : NP\_00046 ; *Bos taurus* : P23795 ; *Felix catus* : O6763 ; *Oryctolagus cuniculus* : Q29499 ; *Rattus norvegicus* : P36136 ; *Mus musculus* : P21836 ; *Gallus gallus* : CAC37792 ; *Danio reno* : Q9DDE3 ; *Electrophorus electricus* : 6730113 ; *Torpedo marmorata* : P07692 ; *Torpedo californica* : P04058 ; *Bungarus fasciatus* : Q92035 ; *Myxine glutinosa* : Q92081.

- Cephalocordés : *Branchiostoma floridae* : O76998 et 076999 ; *Branchiostoma lanceolatum* : Q95000 et Q95001.

- Urocordés : *Ciona intestinalis* : SEQ ID NO : 51 ; *Ciona savignyi* : SEQ ID NO : 52.

- Nématodes : *Caenorhabditis elegans* (1 à 4) : P38433, O61371, O61459 et 061372 ; *Caenorhabditis briggsae* (1 à 4) Q27459, O61378Q9NDG9 et Q9NDG8 ; *Dictyocaulus viviparus* : Q9GPLO.

- Insectes : *Anopheles gambiae* (1 et 2) : SEQ ID NO:3 et SEQ ID NO: 53 ; *Aedes aegypti* (1 et 2) : SEQ ID NO: 9 et AAB3500 ; *An. stephensi* : P56161 ;

*Culex pipiens* : SEQ ID NO: 7 (*ace-1*) et Esther data base pour *ace-2* ; *Drosophila melanogaster* : P07140 ; *Lucilia cuprina* : P91954 ; *Musca domestica* : AAK69132.1 ; *Leptinotarsa decemlineata* : Q27677 ; *Apis mellifera* : AAG43568 ; *Nephotettix cincticeps* : AF145235\_1 ; *Schizaphis graminum* : Q9BMJ1.

5                   - Arachnides : *Rhipicephalus appendiculatus* : O62563 ; *Boophilus microplus* (1 et 2) : O45210 et O61864 ; *Boophilus decoloratus* : O61987 ;

                  - Mollusques : *Loligo opalescens* : O97110.

g) Clonage du fragment K et génotypage d'*ace-1* chez *Culex pipiens*

                  L'ADN de moustique a été extrait comme décrit dans Rogers et al.,  
10   [*Plant Molecular Biology manual*, 1988, eds. Gelvin, S.B. et Schilperoot, R.A. (Kluwer Academic Publishers, Boston), VolA6, p1-10]. Les oligonucléotides PkdirAGSG (5'-ATMGWGTTYGAGTACACSGAYTGG-3', SEQ ID NO: 39) et PkrevAGSG (5'-GGCAAARTTKGWCCAGTATCKCAT-3', SEQ ID NO: 40) amplifient un fragment  
15   de 320 pb (fragment K) à partir de l'ADN génomique de plusieurs moustiques. 30 cycles d'amplification PCR ont été réalisés dans les conditions suivantes : 94°C pendant 30s, 50°C pendant 30s à et 72°C pendant 30s. Les séquences ont été déterminées directement sur les produits PCR sur un séquenceur ABI prism 310, à l'aide du kit *Big Dye Terminator*.

                  Le génotypage d'*ace-1* chez *Culex* est réalisé dans les conditions  
20   suivantes : Les fragments K obtenus comme décrit précédemment sont digérés par *EcoRI* et le produit de digestion est séparé par électrophorèse sur un gel d'agarose à 2%. Les profils de restriction montrent : 1 bande (320 pb) chez les moustiques homozygotes résistants RR, 2 bandes (106 pb et 214 pb) chez les moustiques homozygotes SS et 3 bandes (103 pb, 214 pb et 320 pb) chez les moustiques hétérozygotes RS.

25   h) Clonage de l'ADNc d'*ace-1* chez les individus sensibles et résistants

                  L'ADNc du gène *ace-1* de *Culex pipiens* a été obtenu à partir de l'ARN extrait d'individus de la souche sensible de référence S-LAB, au tout premier stade de développement larvaire L1. La transcription inverse a été réalisée avec un oligonucléotide 18T et la SuperScriptII RNaseH (IN VITROGEN), selon les conditions recommandées par le fabricant.  
30

Deux fragments d'ADNc ont été amplifiés par PCR à l'aide d'oligo-nucléotides dégénérés obtenus à partir de l'alignement des séquences des gènes *ace-1* d'*Anopheles Gambiae* et de *Schizaphis graminum* :

- un fragment b (193pb) a été amplifié à l'aide du couple d'amorces PbdirAGSG (5'GGYGCKACMATGTGGAAYCC3', SEQ ID NO: 41) et PbrevAGSG (5'AGCAMRATCACGTTYTCYTCCGAC3', SEQ ID NO: 42).

- un fragment k (320pb) a été amplifié à l'aide du couple d'amorces PkdirAGSG (5'ATMGWGTTYGAGTACACSGAYTGG3', SEQ ID NO: 39) et PkrevAGSG (5'GGCAAARTTKGWCCAGTATCKCAT3', SEQ ID NO: 40).

Les fragments b et k ainsi obtenus ont ensuite été clonés et séquencés, selon les techniques classiques connues en elles-mêmes de l'Homme du métier, telles que décrites dans *Current Protocols in Molecular Biology* (Frederick M. AUSUBEL, 2000, Wiley and son Inc, Library of Congress, USA).

Un fragment d'ADNc de plus grande taille a été amplifié par PCR, à l'aide d'amorces spécifiques de *Culex pipiens* déduites des séquences des fragments b et k précédemment obtenues. A savoir :

- un fragment CulexA (1127 pb) a été amplifié par PCR à l'aide du couple d'amorces amorces: culex-bdir1 (5'TACATCAACGTGGTCGTGCCACG3', SEQ ID NO: 43) et culex-krev1 (5'GTCACGGTTGCTGTTCTCGGG3', SEQ ID NO: 44). Le fragment Culex A de 1127 pb ainsi obtenu a ensuite été cloné et séquencé, comme ci-dessus.

Les extrémités des ADNc ont été amplifiées par la technique RACE (*Rapid Amplification of cDNA Ends*), à l'aide d'un kit commercial (du kit *Gene Racer* (IN VITROGEN) selon les conditions indiquées dans le manuel d'utilisation. Ensuite elles ont été clonées puis séquencées, comme ci-dessus.

#### i) Clonage du gène *ace-1* chez les individus sensibles et résistants

La séquence de l'ADN génomique de la souche *A. Gambiae* KISUMU (souche sensible de référence de l'Afrique de l'Ouest) a été obtenue à partir de moustiques homozygotes.

De manière plus précise, l'ADN de moustique a été extrait comme décrit dans Rogers et al., [*Plant Molecular Biology manual*, 1988, eds. Gelvin, S.B. & Schilperoot, R.A. (Kluwer Academic Publishers, Boston), Vol 6, p1-10]. 3 fragments chevauchants (A, B et C) ont été amplifiés dans les conditions suivantes : 94°C



pendant 30s, 50°C pendant 30s à et 72°C pendant 30s (30 cycles), à l'aide d'amorces synthétisées à partir de la séquence du gène *ace-1*. A savoir :

- le fragment A (1130pb) a été amplifié à l'aide du couple d'amorces AG1-Adir (5'CGACGCCACCTTCACA3', SEQ ID NO: 45) et AG1-Arev (5'GATGGCCCCGCTGGAACAGAT3', SEQ ID NO: 46),
- le fragment B (1167pb) a été amplifié à l'aide du couple d'amorces AG1-Bdir (5'GGGTGCGGGACAACATTACAC3', SEQ ID NO: 47) et AG1-Brev (5'CCCCGACCGACGAAGGA3', SEQ ID NO: 48), et
- le fragment C (876pb) a été amplifié à l'aide du couple d'amorces AG1-Cdir (5'AGATGGTGGGCGACTATCAC3', SEQ ID NO: 49) et AG1-Crev (5'CTCGTCCGCCACCACTTGTT3', SEQ ID NO: 50).

Les séquences des fragments A, B et C ont été déterminées directement sur les produits PCR, à l'aide d'oligonucléotides internes, inclus dans ces fragments, en utilisant le kit *Big Dye Terminator* et un séquenceur ABI prism 310.

#### 15 **EXEMPLE 2 : Mise en évidence de 2 gènes *ace* chez *Anopheles gambiae***

Des gènes homologues des gènes d'acétylcholinestérases humaines et de drosophiles ont été recherchés à partir de fragments de séquences déposées dans les bases de données, en utilisant le logiciel TBLASTN. Deux groupes de fragments distincts codant pour une AChE très similaire à celle de la drosophile ont été identifiés. Deux gènes de respectivement 6975 pb (*ace-1*) et 5195 pb (*ace-2*) ont été reconstruits à partir de fragments chevauchants de chaque groupe. L'analyse des gènes à l'aide des logiciels FGESH et BLASTX montre que les gènes *ace-1* et *ace-2* sont constitués respectivement d'au moins 7 et 8 exons codant pour des protéines d'environ 534 et 569 acides aminés, dénommées respectivement AChE1 et AChE2. Toutefois, cette analyse n'a pas permis de déterminer avec certitude la séquence des extrémités 5' et 3' de l'ADNc et les séquences NH2 et COOH des protéines correspondantes, qui ne sont pas conservées entre les différentes AChE.

L'analyse des séquences en acides aminés confirme que les protéines AChE1 et d'AChE2 sont très homologues à l'AChE de *Drosophila* (BLASTP :  $P < e^{-180}$ ) et contiennent un motif canonique FGESAG autour de la sérine en position 200, en référence à la séquence de l'Ache de *Torpedo* (S<sub>200</sub> figure 1), qui est caractéristique du site actif des AChE. En outre d'autres motifs caractéristiques des AChE ont égale-

ment été retrouvés dans les deux séquences (AChE1 et AChE2): le site de liaison à la choline (résidu Tryptophane en position 84, W84), les trois résidus de la triade catalytique (résidus sérine, acide glutamique et histidine, respectivement en positions 200, 327 et 440 : S<sub>200</sub>, E<sub>327</sub> et H<sub>440</sub>), les six résidus cystéine potentiellement impliqués dans  
 5 des ponts disulfures conservés (C<sub>67</sub>-C<sub>94</sub>; C<sub>254</sub>-C<sub>265</sub>; C<sub>402</sub>-C<sub>521</sub>), et des résidus aromati-  
 ques bordant la gorge du site actif (10 et 11 résidus , respectivement pour AChE1 et AChE2).

Dans les deux séquences, on observe la présence d'un résidu phénylalanine en position 290 (F290) mais pas en position 288 ; cette caractéristique  
 10 commune aux AChE d'invertébrés est responsable d'une plus large spécificité de substrat des AChE d'invertébrés, par rapport à celles de vertébrés.

L'analyse des séquences C-terminales des AChE de diptère montre la présence d'un peptide hydrophobe correspondant à un signal d'addition d'un glycolipide, indiquant le clivage post-traductionnel d'un fragment C-terminal et l'addition  
 15 d'une résidu d'ancrage glycolipidique comme chez *Drosophila* et d'autres espèces de moustiques. Dans toutes les séquences on observe également la présence d'un résidu cystéine dans la séquence C-terminale précédant le site potentiel de clivage du peptide hydrophobe. Ce résidu cystéine pourrait être impliqué dans une liaison disulfure intermoléculaire, liant les deux sous-unités catalytiques du dimère d'AChE.

Les protéines AChE1 et AChE2 d'*An. gambiae* présentent 53 % de similarité entre elles et montrent respectivement : 76 % et 55 % de similarité avec l'AChE de *Schizaphis graminum* (numéro d'accension NCBI AAK09373 ou GENBANK 12958609), 53 % et 98 % de similarité avec l'AChE d'*An. stephensi* (GENBANK 2494391), 54 % et 95 % de similarité avec l'AChE d'*Aedes aegypti*  
 25 (GENBANK 2133626), 52 % et 83 % de similarité avec l'AChE de *Drosophila* (GENBANK 17136862).

La différence majeure entre AChE1 et AChE2 réside dans une insertion de 31 acides aminés dans la séquence d'AChE2 (figure 1). Cette séquence, dénommée "insertion hydrophilique" dans l'AChE de *Drosophila*, est absente dans les  
 30 AChEs de vertébrés et de nématodes et pourrait être caractéristique de l'AChE2, au moins chez les diptères.

Ces résultats démontrent la présence de deux gènes *ace* dans le génome d'*Anopheles gambiae*, l'un codant pour AChE1 qui est apparentée à l'AChE de *Schizaphis graminum*, et l'autre pour AChE2 qui est apparentée à l'AChE de *Drosophila* et aux AChEs connues de moustiques. La présence d'autres gènes *ace* chez *An. gambiae* est très improbable dans la mesure où des recherches complémentaires dans les bases de données du génome d'*An gambiae*, en utilisant des paramètres moins stringents, ont détecté uniquement des séquences codant pour des alpha-estérases (EC 3.1.1) et des carboxylestérases (EC 3.1.1.1).

**EXEMPLE 3 : Mise en évidence d'un unique gène *ace* chez *Drosophila melanogaster***

La présence d'un gène homologue du gène *ace-1* a été recherchée dans le génome de *Drosophila*. Les recherches TBLASTN ont permis de détecter le gène *ace* précédemment identifié, homologue du gène *ace-2* d'*Anopheles gambiae* mais n'ont pas permis de détecter d'autres séquences homologues du gène *ace-1*. Des recherches à l'aide de paramètres moins stringents ont permis de détecter uniquement des alpha et des carboxylestérases. Ces résultats démontrent que le génome de la drosophile contient un unique gène *ace* (*ace-2*).

**EXEMPLE 4 : Mise en évidence d'au moins deux gènes *ace* chez les autres espèces de moustiques**

La présence du gène *ace-1* dans le génome d'autres espèces de moustiques a été analysée par PCR à l'aide d'oligonucléotides dégénérés (PdirAGSG et PrevAGSG, SEQ ID NO: 39 et 40) permettant d'amplifier un fragment exonique (fragment K, d'environ 320 pb figure 1), correspondant à des séquences conservées entre les séquences d'AChE1 d'*An. gambiae* et *Schizaphis graminum* mais divergentes entre les séquences d'AChE1 et AChE2 d'*An. gambiae*.

La séquence des produits PCR obtenus à partir de l'ADN génomique de différentes espèces de moustiques, montre un pourcentage d'identité très élevé entre les séquences d'*Anopheles*, *Culex* et *Aedes*. En outre, la plupart des substitutions sont silencieuses puisque les séquences en acides aminés déduites de ces séquences nucléotidiques ne diffèrent entre elles que part 5 à 6 acides aminés (Figure 2A). Le fragment K a également été amplifié par RT-PCR à partir de l'ARNm de *C. pipiens*, indiquant que le gène *ace-1* est exprimé sous forme d'ARNm ; ce résultat est en

accord avec l'existence, chez *C. pipiens*, de deux AChEs possédant des propriétés catalytiques distinctes.

**EXEMPLE 5 : Analyse de la liaison entre le gène *ace-1* et la résistance aux insecticides**

5 Afin d'analyser la liaison entre le gène *ace-1* et la résistance aux insecticides, le fragment K amplifié à partir de l'ADN génomique de *C. pipiens* résistants (souche R), a été séquencé. La comparaison des séquences du fragment K entre les souches S et R montre des différences au niveau de 3 nucléotides (substitutions silencieuses, Figure 2B). L'une de ces substitutions affecte un site *EcoRI*, ce qui  
10 permet de différencier facilement le locus *ace-1* des souches S et R par PCR-RFLP : les profils de restriction montrent 1 bande (320 pb) chez les individus homozygotes résistants, 2 bandes (106 pb et 214 pb) chez les moustiques homozygotes SS et 3 bandes (103 pb, 214 pb et 320 pb) chez les moustiques hétérozygotes RS (figure 2C).

La liaison entre le gène *ace-1* et la résistance au propoxur a été étudiée, en triple, de la façon suivante : des larves de croisement en retour (S x R) x S ont  
15 été traitées par une dose létale pour les individus sensibles et le génotype d'*ace-1* a été analysé chez les survivants, par PCR-RFLP.

Les résultats montrent que l'exposition au propoxur tue 50 % des larves dans tous les croisements en retour, c'est à dire tous les individus sensibles.  
20 Toutes les larves survivantes (100 pour chaque croisement en retour, 300 au total) montrent un profil hétérozygote en RFLP, indiquant qu'elles possèdent toutes une copie du gène *ace-1* de la souche R.

Ces résultats démontrent que le gène *ace-1* est lié de façon très étroite avec la résistance aux insecticides (moins de 1 % de recombinaison avec un  
25 degré de confiance de 0,05).

**EXEMPLE 6 : Analyse de la phylogénie des gènes *ace-1* et *ace-2*.**

Des arbres phylogénétiques ont été construits à partir des séquences des régions conservées des AChE d'*An gambiae* (SEQ ID NO: 1 et fragment 34-393 de la séquence SEQ ID NO: 53, figure 1), des fragments K de *C. pipiens* et *Aedes aegypti* (SEQ ID NO: 8 et 9) et de 33 séquences d'AChE disponibles dans  
30 GENBANK, à l'aide de la méthode du plus proche voisin (*neighbour-joining method*), comme décrit dans le matériels et méthodes.

La figure 3 illustre l'hétérogénéité du nombre de gènes *ace* au cours de l'évolution du règne animal. Chez les cordés, les céphalocordés possèdent au moins deux gènes *ace* alors que les urocordés n'en possèdent qu'un seul, comme déduit de l'analyse de leur génome. Chez les arthropodes, les diptères possèdent, soit un seul gène *ace* (*Drosophila* du sous-ordre des brachycères) ou deux gènes *ace* (moustiques du sous-ordre des nématocères). La topologie de l'arbre montre que ces deux gènes *ace* se sont dupliqués très précocement au cours de l'évolution, probablement avant la séparation entre les protostomes et les deutérostomes. Ces résultats sont supportés par le fait que les AChE de mollusques, de nématodes et d'arthropodes se ramifient à partir des séquences des cordés (*craniatia*, céphalocordés et urocordés). Les résultats montrent que les arthropodes et les nématodes possèdent une AChE apparentée.

Ces résultats indiquent que les gènes *ace-1* et *ace-2* identifiés chez les insectes proviennent d'un événement de duplication très ancien et que l'absence du gène *ace-1*, au moins chez certaines espèces du sous-ordre des brachycères (*Drosophila*) résulte de la perte d'un gène *ace* plutôt que d'une duplication récente du gène *ace* chez les nématocères. Ces résultats suggèrent également que les extrapolations faites à partir d'études chez *D. melanogaster* sont à considérer avec réserve dans la mesure où la situation de *Drosophila* n'est ni représentative des diptères ni de l'ensemble de la classe des insectes.

#### 20 **EXEMPLE 7 : Détermination de la séquence complète de l'ADNc d'*ace-1***

L'ADNc d'*ace-1* a été cloné à partir d'une souche d'*Anopheles gambiae* de l'Afrique de l'Ouest (souche KISUMU) et d'une souche de *Culex pipiens* (souche S-LAB) sensibles aux insecticides de la classe des organophosphorés et des carbamates, comme décrit dans le matériels et méthodes

25 La séquence complète de l'ADNc correspond respectivement aux séquences SEQ ID NO: 4 et SEQ ID NO: 6 et qui codent pour une protéine de respectivement 643 et 702 acides aminés (SEQ ID NO:5 et SEQ ID NO: 7).

#### **EXEMPLE 8 : Détermination de la séquence complète du gène *ace-1***

30 La séquence complète du gène *ace-1* a été déterminée à partir de l'ADN génomique d'une souche d'*Anopheles gambiae* de l'Afrique de l'Ouest (souche KISUMU), comme décrit dans le matériels et méthodes.

La séquence complète du gène *ace-1* correspond à la séquence SEQ ID NO: 23 qui présente une organisation intron-exon comprenant au moins 7 exons (Tableau I).

5 Ainsi que cela ressort de ce qui précède, l'invention ne se limite nullement à ceux de ses modes de mise en œuvre, de réalisation et d'application qui ~~viennent d'être décrits de façon plus explicite ;~~ elle en embrasse au contraire toutes les variantes qui peuvent venir à l'esprit du technicien en la matière, sans s'écarter du cadre, ni de la portée, de la présente invention.

### REVENDEICATIONS

1°) Acétylcholinestérase d'insecte, caractérisée en ce qu'elle comprend une région catalytique centrale qui présente une séquence en en acides aminés sélectionnée dans le groupe constitué par la séquence SEQ ID NO: 1 et les  
5 séquences présentant au moins 60 % d'identité ou 70 % de similarité avec la séquence SEQ ID NO: 1, à l'exclusion de l'acétylcholinestérase de séquence NCBI AAK0973.

2°) Acétylcholinestérase selon la revendication 1, caractérisée en ce qu'elle correspond à celle d'un insecte de la famille des *Culicidae*, choisi parmi les genres *Culex*, *Aedes* et *Anopheles*.

10 3°) Acétylcholinestérase selon la revendication 2, caractérisée en ce qu'elle est constituée par une séquence sélectionnée dans le groupe constitué par les séquences SEQ ID NO: 3, SEQ ID NO: 5 et SEQ ID NO: 7.

15 4°) Acétylcholinestérase selon la revendication 2, caractérisée en ce que ladite région catalytique centrale comprend une séquence sélectionnée dans le groupe constitué par les séquences SEQ ID NO: 8 à 21.

5°) Peptide, caractérisé en ce qu'il est constitué par un fragment d'au moins 7 acides aminés de l'acétylcholinestérase selon l'une quelconque des revendications 1 à 4.

20 6°) Molécule d'acide nucléique isolée, caractérisée en ce qu'elle présente une séquence sélectionnée dans le groupe constitué par :

- les séquences codant pour une acétylcholinestérase selon l'une quelconque des revendications 1 à 4 (ADNc et gène *ace-1*);

- les séquences complémentaires des séquences précédentes, sens ou anti-sens, et

25 - les fragments d'au moins 8 pb, de préférence de 15 pb à 500 pb des séquences précédentes.

7°) Molécule d'acide nucléique selon la revendication 6, caractérisée en ce qu'elle est sélectionnée dans le groupe constitué par les ADNc de séquence SEQ ID NO: 2, SEQ ID NO:4 et SEQ ID NO: 6 et les ADN génomiques de séquence SEQ  
30 ID NO: 22 et SEQ ID NO: 23.

8°) Molécule d'acide nucléique selon la revendication 6, caractérisée en ce qu'elle est sélectionnée dans le groupe constitué par les amorces de séquence SEQ ID NO: 39 à 50 et les fragments de séquence SEQ ID NO: 24 à 38.

5 9°) Méthode de détection d'insectes porteurs d'une résistance aux insecticides de la classe des organophosphorés et des carbamates, caractérisée en ce qu'elle comprend:

---

- la préparation d'un échantillon d'acides nucléiques à partir d'insectes à tester, et

10 - la détection par tout moyen approprié, de la présence dans ledit échantillon d'acides nucléiques, d'une mutation dans le gène *ace-1* tel que défini à la revendication 6 ou à la revendication 7.

10°) Méthode selon la revendication 9, caractérisée en ce que ladite détection comprend :

15 - l'amplification d'un fragment d'environ 320 pb à l'aide du couple d'amorces SEQ ID NO: 39 et 40,

- la digestion dudit fragment à l'aide d'une enzyme de restriction appropriée, et

- l'analyse du profil de restriction obtenu.

20 11°) Méthode selon la revendication 10, caractérisée en ce que ladite enzyme de restriction est *EcoRI*.

12°) Vecteur recombinant, caractérisé en ce qu'il comprend un insert sélectionné dans le groupe constitué par les molécules d'acides nucléiques selon l'une quelconque des revendications 6 à 8.

25 13°) Cellules, caractérisées en ce qu'elles sont modifiées par un vecteur recombinant selon la revendication 12.

14°) Anticorps, caractérisés en ce qu'ils sont dirigés contre l'acétylcholinestérase selon l'une quelconque des revendications 1 à 4 ou le peptide selon la revendication 5.

30 15°) Réactif de détection d'insectes porteurs d'une résistance aux insecticides de la classe des organophosphorés et des carbamates, caractérisé en ce qu'il est sélectionné dans le groupe constitué par les molécules d'acides nucléiques et



leurs fragments selon l'une quelconque des revendications 6 à 8 et les anticorps selon la revendication 14.

16°) Animal invertébré transgénique, caractérisé en ce qu'il contient des cellules transformées par au moins une molécule d'acide nucléique selon la revendication 6 ou la revendication 7.

17°) Méthode de criblage d'une substance insecticide, caractérisée en ce qu'elle comprend :

a) la mise en contact de la substance à tester avec une acétylcholinestérase selon l'une quelconque des revendications 1 à 4, un extrait de cellules modifiées telles que définies à la revendication 13 ou un échantillon biologique d'un animal transgénique tel que défini à la revendication 16, en présence d'acétylcholine ou de l'un de ses dérivés, et

b) la mesure par tout moyen approprié, de l'activité acétylcholinestérase du mélange obtenu en a), et

c) la sélection des substances capables d'inhiber ladite activité.

18°) Méthode de criblage de substances insecticides, caractérisée en ce qu'elle comprend :

- la mise en contact d'une substance à tester avec un animal transgénique selon la revendication 16, et

- la mesure de la survie de l'animal.

19°) Réactif de criblage de substances insecticides, caractérisé en ce qu'il est sélectionné dans le groupe constitué par les acétylcholinestérases selon l'une quelconque des revendications 1 à 4, les vecteurs recombinants selon la revendication 12, les cellules modifiées selon la revendication 13 et les animaux transgéniques selon la revendication 16.

20°) Trousse de détection et/ou de criblage, caractérisée en ce qu'elle inclut au moins un réactif selon la revendication 15 ou la revendication 19.

67 70 84 92

**Ace1**  
Agam1 DPLVNTDFGRIGIT/DAPSGLKVD/WLGIPIYAKPPVGLRFRHPRPAEKNTG-----VLNTTTPNSCVQIVDTVEGDEPGATMNNPFLSEDCLYL  
Sgra DPLIHTNKGKIRGITQTATTGKLVDAHLGIPYAKKPIGLRFRHPRPTDMDCTTTPETILNCTTPNTQVQIFDTLEGDEPGATMNNPFLSEDCLYL

**Ace2**  
Agam2 DPLVNTTSGLIRGSTMTVGGGVEVIVNGVPFAKPPVDSLREKPEPAEEMG-----VLNTTTPNSCVQIVDTVEGDEPGATMNNPFLSEDCLYL  
Aste DPLVNTTSGLIRGSTMTVGGGVEVIVNGVPFAKPPVDSLREKPEPAEEMG-----VLNTTTPNSCVQIVDTVEGDEPGATMNNPFLSEDCLYL  
Aaeq2 DPLVNTTSGLIRGSTMTVGGGVEVIVNGVPFAKPPVDSLREKPEPAEEMG-----VLNTTTPNSCVQIVDTVEGDEPGATMNNPFLSEDCLYL  
Cmel DPLVNTTSGLIRGSTMTVGGGVEVIVNGVPFAKPPVDSLREKPEPAEEMG-----VLNTTTPNSCVQIVDTVEGDEPGATMNNPFLSEDCLYL  
Lcup DPLVNTTSGLIRGSTMTVGGGVEVIVNGVPFAKPPVDSLREKPEPAEEMG-----VLNTTTPNSCVQIVDTVEGDEPGATMNNPFLSEDCLYL  
Mdom DPLVNTTSGLIRGSTMTVGGGVEVIVNGVPFAKPPVDSLREKPEPAEEMG-----VLNTTTPNSCVQIVDTVEGDEPGATMNNPFLSEDCLYL  
Cpip2 DPLVNTTSGLIRGSTMTVGGGVEVIVNGVPFAKPPVDSLREKPEPAEEMG-----VLNTTTPNSCVQIVDTVEGDEPGATMNNPFLSEDCLYL

114 121 130

**Ace1**  
Agam1 NVVAER--PRPK-----NAAVMLWIFGGGFYSGTATLDVYDHRALASEENVIVSVLQYRVASLGELFLG-----  
Sgra NVVAER--PRPK-----NAAVMLWIFGGGFYSGTATLDVYDHRALASEENVIVSVLQYRVASLGELFLG-----

**Ace2**  
Agam2 NIIVPTTIRHREGGLFENNDYFQDDDFQRCQCSK-----GHAMVWVINGGGFSGTSTLDIYAEILAAVGNVIVSMQYRVASLGELYLAFLYING  
Aste NIIVPTTIRHREGGLFENNDYFQDDDFQRCQCSK-----GHAMVWVINGGGFSGTSTLDIYAEILAAVGNVIVSMQYRVASLGELYLAFLYING  
Aaeq2 NIIVPTTIRHREGGLFENNDYFQDDDFQRCQCSK-----GHAMVWVINGGGFSGTSTLDIYAEILAAVGNVIVSMQYRVASLGELYLAFLYING  
Cmel NVVAER--PRPK-----NAAVMLWIFGGGFYSGTATLDVYDHRALASEENVIVSVLQYRVASLGELFLG-----  
Lcup NIIVPTTIRHREGGLFENNDYFQDDDFQRCQCSK-----GHAMVWVINGGGFSGTSTLDIYAEILAAVGNVIVSMQYRVASLGELYLAFLYING  
Mdom NIIVPTTIRHREGGLFENNDYFQDDDFQRCQCSK-----GHAMVWVINGGGFSGTSTLDIYAEILAAVGNVIVSMQYRVASLGELYLAFLYING  
Cpip2 NIIVPTTIRHREGGLFENNDYFQDDDFQRCQCSK-----GHAMVWVINGGGFSGTSTLDIYAEILAAVGNVIVSMQYRVASLGELYLAFLYING

200 233 254

**Ace1**  
Agam1 --TPEAFGNAGLFDONLALRWVRCNIHREGGDESRVTLLEGESAGAVSVSLHLLSALSRLDFCRAILQSGSPAPALVALVSREATLRLALDASAVGCPHEP  
Sgra --TPEAFGNAGLFDONLALRWVRCNIHREGGDESRVTLLEGESAGAVSVSLHLLSALSRLDFCRAILQSGSPAPALVALVSREATLRLALDASAVGCPHEP

**Ace2**  
Agam2 -YEEDAPGNMGMDQALAIRWENKAEFGGDPOLITTEGESAGGSSVSLHLLSPVTRGLSKRILQSGILAPPSHTAEALQIAEGGIDECNLT  
Aste -YEEDAPGNMGMDQALAIRWENKAEFGGDPOLITTEGESAGGSSVSLHLLSPVTRGLSKRILQSGILAPPSHTAEALQIAEGGIDECNLT  
Aaeq2 -YEEDAPGNMGMDQALAIRWENKAEFGGDPOLITTEGESAGGSSVSLHLLSPVTRGLSKRILQSGILAPPSHTAEALQIAEGGIDECNLT  
Cmel -YEEDAPGNMGMDQALAIRWENKAEFGGDPOLITTEGESAGGSSVSLHLLSPVTRGLSKRILQSGILAPPSHTAEALQIAEGGIDECNLT  
Lcup -YEEDAPGNMGMDQALAIRWENKAEFGGDPOLITTEGESAGGSSVSLHLLSPVTRGLSKRILQSGILAPPSHTAEALQIAEGGIDECNLT  
Mdom -YEEDAPGNMGMDQALAIRWENKAEFGGDPOLITTEGESAGGSSVSLHLLSPVTRGLSKRILQSGILAPPSHTAEALQIAEGGIDECNLT  
Cpip2 -YEEDAPGNMGMDQALAIRWENKAEFGGDPOLITTEGESAGGSSVSLHLLSPVTRGLSKRILQSGILAPPSHTAEALQIAEGGIDECNLT

265 279 288 290 327 330 331 334

**Ace1**  
Agam1 --SKLSDAVECLGKDPHVLVNNWGTG-LGICSEFFVEVVDGAELODETQORSLASGRFKKTELTGNTSEGYFIIYVLTLLRKEZGVTVTRFELQAA  
Sgra --NTIHKTVZGLKANSVVMVEKE-DHVAICFSEFFVEVVDGAELODDHQKSLSTNNFKXTNLMGNSNEEGYISIFVLTLLFKKEENVMSVSENFIXA

**Ace2**  
Agam2 LKESSTVMOCRRNVDAKTSVQCVNSYSGLIGESAPTTDGVEMTADENTMLREANLEGIDILVGSNR EGYFFLYIDYDFDKCAAGSLRORFLYET  
Aste LKESSTVMOCRRNVDAKTSVQCVNSYSGLIGESAPTTDGVEMTADENTMLREANLEGIDILVGSNR EGYFFLYIDYDFDKCAAGSLRORFLYET  
Aaeq2 LKESSTVMOCRRNVDAKTSVQCVNSYSGLIGESAPTTDGVEMTADENTMLREANLEGIDILVGSNR EGYFFLYIDYDFDKCAAGSLRORFLYET  
Cmel LKESSTVMOCRRNVDAKTSVQCVNSYSGLIGESAPTTDGVEMTADENTMLREANLEGIDILVGSNR EGYFFLYIDYDFDKCAAGSLRORFLYET  
Lcup LKESSTVMOCRRNVDAKTSVQCVNSYSGLIGESAPTTDGVEMTADENTMLREANLEGIDILVGSNR EGYFFLYIDYDFDKCAAGSLRORFLYET  
Mdom LKESSTVMOCRRNVDAKTSVQCVNSYSGLIGESAPTTDGVEMTADENTMLREANLEGIDILVGSNR EGYFFLYIDYDFDKCAAGSLRORFLYET  
Cpip2 LKESSTVMOCRRNVDAKTSVQCVNSYSGLIGESAPTTDGVEMTADENTMLREANLEGIDILVGSNR EGYFFLYIDYDFDKCAAGSLRORFLYET

PdrAGSG ← K fragment 402 432 440 442

**Ace1**  
Agam1 VRELNPVYNGAARQAFVEFYTDWTEPDNPNSNRDALDKMVGDYHTCNVNEFAQRYAEZGNVVMYLYTHRSKGNPFRITGMGDEINVEGEELNSD  
Sgra IQQLNPADAAXVSAIEFEYTDWTEPDNPNSNRDALDKMVGDYHTCNVNEFAQRYAEZGNVVMYLYTHRSKGNPFRITGMGDEINVEGEELNSD  
Cpip1 IEQQLNPADAAXVSAIEFEYTDWTEPDNPNSNRDALDKMVGDYHTCNVNEFAQRYAEZGNVVMYLYTHRSKGNPFRITGMGDEINVEGEELNSD  
Aaeq1 TEPENPNSNRDALDKMVGDYHTCNVNEFAQRYAEZGNVVMYLYTHRSKGNPFRITGMGDEINVEGEELNSD

**Ace2**  
Agam2 MNTIENASEPPEPAITFEYTSWE-NDYONKHVGRVGDHEICPENEFALGLTERGASVHYFFTHRTSTSLWENGVHGGDEVYIEGEMNAN  
Aste MNTIENASEPPEPAITFEYTSWE-NDYONKHVGRVGDHEICPENEFALGLTERGASVHYFFTHRTSTSLWENGVHGGDEVYIEGEMNAN  
Aaeq2 MNTIENASEPPEPAITFEYTSWE-NDYONKHVGRVGDHEICPENEFALGLTERGASVHYFFTHRTSTSLWENGVHGGDEVYIEGEMNAN  
Cmel MNTIENASEPPEPAITFEYTSWE-NDYONKHVGRVGDHEICPENEFALGLTERGASVHYFFTHRTSTSLWENGVHGGDEVYIEGEMNAN  
Lcup MNTIENASEPPEPAITFEYTSWE-NDYONKHVGRVGDHEICPENEFALGLTERGASVHYFFTHRTSTSLWENGVHGGDEVYIEGEMNAN  
Mdom MNTIENASEPPEPAITFEYTSWE-NDYONKHVGRVGDHEICPENEFALGLTERGASVHYFFTHRTSTSLWENGVHGGDEVYIEGEMNAN  
Cpip2 MNTIENASEPPEPAITFEYTSWE-NDYONKHVGRVGDHEICPENEFALGLTERGASVHYFFTHRTSTSLWENGVHGGDEVYIEGEMNAN

→ PrevAGSG 521

**Ace1**  
Agam1 LGYTEDEKDFSPKIMRYVNSFKTQNNPNTASSEF-PEWPKHTAHGRHYLELG-----LNTSFVGRGERLRQCAEKKYLQOLVAA  
Sgra KRYEISEIELSKMMRYWNTNEAKTGNBSKTLEGSWVTPKNEVHTAYGKEFLD-----TNNTSIGVGRERLEQCAEKKYVLDLTAI  
Cpip1 LGYTEDEKDFSPKIMRYVNSFKTQNNPNTASSEF-PEWPKHTAHGRHYLELG-----LNTSFVGRGERLRQCAEKKYLQOLVAA  
Aaeq1 LGYTEDEKDFSPKIMRYVNSFKTQNNPNTASSEF-PEWPKHTAHGRHYLELG-----LNTSFVGRGERLRQCAEKKYLQOLVAA

**Ace2**  
Agam2 LQYQREEDLSRPMVLSSEFARTGNFLEGE-----HWELYTRNFIYFI-NAEGEDLRGKYGRGEMTSGAEVYFLERLAA  
Aste LQYQREEDLSRPMVLSSEFARTGNFLEGE-----HWELYTRNFIYFI-NAEGEDLRGKYGRGEMTSGAEVYFLERLAA  
Aaeq2 LQYQREEDLSRPMVLSSEFARTGNFLEGE-----HWELYTRNFIYFI-NAEGEDLRGKYGRGEMTSGAEVYFLERLAA  
Cmel LQYQREEDLSRPMVLSSEFARTGNFLEGE-----HWELYTRNFIYFI-NAEGEDLRGKYGRGEMTSGAEVYFLERLAA  
Lcup LQYQREEDLSRPMVLSSEFARTGNFLEGE-----HWELYTRNFIYFI-NAEGEDLRGKYGRGEMTSGAEVYFLERLAA  
Mdom LQYQREEDLSRPMVLSSEFARTGNFLEGE-----HWELYTRNFIYFI-NAEGEDLRGKYGRGEMTSGAEVYFLERLAA  
Cpip2 LQYQREEDLSRPMVLSSEFARTGNFLEGE-----HWELYTRNFIYFI-NAEGEDLRGKYGRGEMTSGAEVYFLERLAA

FIGURE 1

	1								80
Ae alb	TEPDNPNSNR	DALDKMVG DY	HFTCNVNEFA	QRYAEEGNNV	YMYLYTHRSK	GNPWPRTG V	MHGDEINYV F	GEPLNSDLG Y	
Ae aeg	TEPDNPNSNR	DALDKMVG DY	HFTCNVNEFA	QRYAEEGNNV	YMYLYTHRSK	GNPWPRTG V	MHGDEINYV F	GEPLNSDLG Y	
An alb	TEPDNPNSNR	DALDKMVG DY	HFTCNVNEFA	QRYAEEGNNV	YMYLYTHRSK	GNPWPRTG V	MHGDEINYV F	GEPLNP TLG Y	
An gam	TEPDNPNSNR	DALDKMVG DY	HFTCNVNEFA	QRYAEEGNNV	YMYLYTHRSK	GNPWPRTG V	MHGDEINYV F	GEPLNP TLG Y	
An fun	TEPDNPNSNR	DALDKMVG DY	HFTCNVNEFA	QRYAEEGNNV	YMYLYTHRSK	GNPWPRTG V	MHGDEINYV F	GEPLNP TLG Y	
An nil	TEPDNPNSNR	DALDKMVG DY	HFTCNVNEFA	QRYAEEGNNV	YMYLYTHRSK	GNPWPRTG V	MHGDEINYV F	GEPLNP TLG Y	
An sac	TEPDNPNSNR	DALDKMVG DY	HFTCNVNEFA	QRYAEEGNNV	YMYLYTHRSK	GNPWPRTG V	MHGDEINYV F	GEPLNP TLG Y	
An pse	TEPDNPNSNR	DALDKMVG DY	HFTCNVNEFA	QRYAEEGNNV	YMYLYTHRSK	GNPWPRTG V	MHGDEINYV F	GEPLNP TLG Y	
Cx Pip	TEPDNPNSNR	DALDKMVG DY	HFTCNVNEFA	QRYAEEGNNV	YMYLYTHRSK	GNPWPRTG V	MHGDEINYV F	GEPLNP TLG Y	

	81	91
Ae alb	<u>MD</u> DEKOF SRK I	
Ae aeg	<u>MD</u> DEKOF SRK I	
An alb	<u>TD</u> DEKOF SRK I	
An gam	<u>TE</u> DEKOF SRK I	
An fun	<u>TE</u> DEKOF SRK I	
An nil	<u>TE</u> DEKOF SRK I	M
An sac	<u>TD</u> DEKOF SRK I	
An pse	<u>TE</u> DEKOF SRK I	
Cx Pip	<u>OD</u> DEKOF SRK I	

Acel-SLAB ATGAACCGGACAACCCGAACAGCAACCGTGACGCGTGGACAAGATGGTCGGGGATTATCACTTCACCTGCAACGTGAA  
 Acel-SR ATGAACCGGACAACCCGAACAGCAACCGTGACGCGTGGACAAGATGGTCGGGGATTATCACTTCACCTGCAACGTGAA  
  
 EcoRI 100 120 140 160  
 Acel-SLAB CGAATTCGCCCGACGGGTACGCCGAGGAGGGCAACAACGTGTTTCATGTACCTGTACACGCACAGAAGCAAAGGAAATCCCT  
 Acel-SR CGAATTCGCCCGACGGGTACGCCGAGGAGGGCAACAATGTGTTTCATGTACCTGTACACGCACAGAAGCAAAGGAAATCCCT  
  
 180 200 220 240  
 Acel-SLAB GGCCGAGGTGGACGGCGTGATGCACGGCGACGAGATCAACTACGTGTTTGGCGAACCGCTGAACTCGGCCCTCGGCTAC  
 Acel-SR GGCCGAGGTGGACGGCGTGATGCACGGCGACGAGATCAACTACGTGTTTGGCGAACCGCTGAACTCGGCCCTCGGCTAC  
  
 260  
 Acel-SLAB CAGGACGACGAGAAGGACTTTAGCCGGAATAATT  
 Acel-SR CAGGACGACGAGAAGGACTTTAGCCGGAATAATT

**S-LAB      SR      Back-cross**

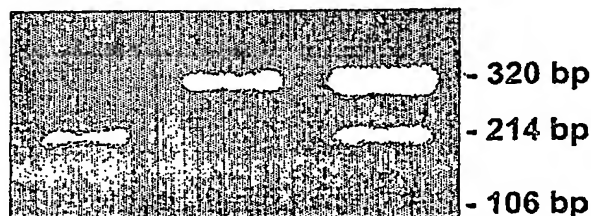


FIGURE 2



**FIGURE 3**

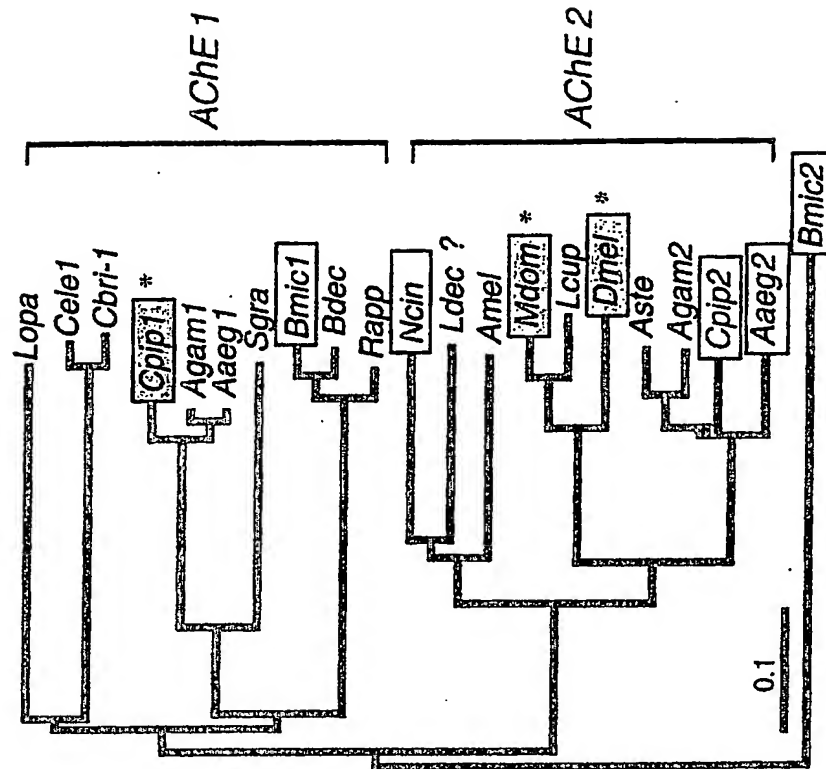


FIGURE 4

## LISTE DE SEQUENCES

<110> CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE -CNRS

<120> Nouveau gène de l'acétylcholinestérase responsable de la  
résistance aux insecticides et ses applications

<130> SF644FR80

<140>

<141>

<160> 53

<170> PatentIn Ver. 2.1

<210> 1

<211> 524

<212> PRT

<213> Anopheles gambiae

<400> 1

Asp Pro Leu Val Val Asn Thr Asp Lys Gly Arg Ile Arg Gly Ile Thr  
1 5 10 15

Val Asp Ala Pro Ser Gly Lys Lys Val Asp Val Trp Leu Gly Ile Pro  
20 25 30

Tyr Ala Gln Pro Pro Val Gly Pro Leu Arg Phe Arg His Pro Arg Pro  
35 40 45

Ala Glu Lys Trp Thr Gly Val Leu Asn Thr Thr Thr Pro Pro Asn Ser  
50 55 60

Cys Val Gln Ile Val Asp Thr Val Phe Gly Asp Phe Pro Gly Ala Thr  
65 70 75 80

Met Trp Asn Pro Asn Thr Pro Leu Ser Glu Asp Cys Leu Tyr Ile Asn  
85 90 95

Val Val Ala Pro Arg Pro Arg Pro Lys Asn Ala Ala Val Met Leu Trp  
100 105 110

Ile Phe Gly Gly Gly Phe Tyr Ser Gly Thr Ala Thr Leu Asp Val Tyr  
115 120 125

Asp His Arg Ala Leu Ala Ser Glu Glu Asn Val Ile Val Val Ser Leu  
130 135 140

Gln Tyr Arg Val Ala Ser Leu Gly Phe Leu Phe Leu Gly Thr Pro Glu  
145 150 155 160

Ala Pro Gly Asn Ala Gly Leu Phe Asp Gln Asn Leu Ala Leu Arg Trp  
165 170 175

Val Arg Asp Asn Ile His Arg Phe Gly Gly Asp Pro Ser Arg Val Thr  
180 185 190

Leu Phe Gly Glu Ser Ala Gly Ala Val Ser Val Ser Leu His Leu Leu  
 195 200 205  
 Ser Ala Leu Ser Arg Asp Leu Phe Gln Arg Ala Ile Leu Gln Ser Gly  
 210 215 220  
 Ser Pro Thr Ala Pro Trp Ala Leu Val Ser Arg Glu Glu Ala Thr Leu  
 225 230 235 240  
 Arg Ala Leu Arg Leu Ala Glu Ala Val Gly Cys Pro His Glu Pro Ser  
 245 250 255  
 Lys Leu Ser Asp Ala Val Glu Cys Leu Arg Gly Lys Asp Pro His Val  
 260 265 270  
 Leu Val Asn Asn Glu Trp Gly Thr Leu Gly Ile Cys Glu Phe Pro Phe  
 275 280 285  
 Val Pro Val Val Asp Gly Ala Phe Leu Asp Glu Thr Pro Gln Arg Ser  
 290 295 300  
 Leu Ala Ser Gly Arg Phe Lys Lys Thr Glu Ile Leu Thr Gly Ser Asn  
 305 310 315 320  
 Thr Glu Glu Gly Tyr Tyr Phe Ile Ile Tyr Tyr Leu Thr Glu Leu Leu  
 325 330 335  
 Arg Lys Glu Glu Gly Val Thr Val Thr Arg Glu Glu Phe Leu Gln Ala  
 340 345 350  
 Val Arg Glu Leu Asn Pro Tyr Val Asn Gly Ala Ala Arg Gln Ala Ile  
 355 360 365  
 Val Phe Glu Tyr Thr Asp Trp Thr Glu Pro Asp Asn Pro Asn Ser Asn  
 370 375 380  
 Arg Asp Ala Leu Asp Lys Met Val Gly Asp Tyr His Phe Thr Cys Asn  
 385 390 395 400  
 Val Asn Glu Phe Ala Gln Arg Tyr Ala Glu Glu Gly Asn Asn Val Tyr  
 405 410 415  
 Met Tyr Leu Tyr Thr His Arg Ser Lys Gly Asn Pro Trp Pro Arg Trp  
 420 425 430  
 Thr Gly Val Met His Gly Asp Glu Ile Asn Tyr Val Phe Gly Glu Pro  
 435 440 445  
 Leu Asn Pro Thr Leu Gly Tyr Thr Glu Asp Glu Lys Asp Phe Ser Arg  
 450 455 460  
 Lys Ile Met Arg Tyr Trp Ser Asn Phe Ala Lys Thr Gly Asn Pro Asn  
 465 470 475 480  
 Pro Asn Thr Ala Ser Ser Glu Phe Pro Glu Trp Pro Lys His Thr Ala  
 485 490 495

His Gly Arg His Tyr Leu Glu Leu Gly Leu Asn Thr Ser Phe Val Gly  
500 505 510

Arg Gly Pro Arg Leu Arg Gln Cys Ala Phe Trp Lys  
515 520

<210> 2

<211> 1932

<212> ADN

<213> Anopheles gambiae

<220>

<221> CDS

<222> (1)..(1932)

<400> 2

atg ttt gtg tgt tgt ttt ttc ttt ctc tct ctc tct ttc tgt ggt tcc	48
Met Phe Val Cys Cys Phe Phe Phe Leu Ser Leu Ser Phe Cys Gly Ser	
1 5 10 15	
aac att tca gac gca ttt ttt aca cca tat ata ggt cac ggt gag tcc	96
Asn Ile Ser Asp Ala Phe Phe Thr Pro Tyr Ile Gly His Gly Glu Ser	
20 25 30	
gta cga att ata gat gcc gag ttg ggc acg ctc gag cat gtc cac agt	144
Val Arg Ile Ile Asp Ala Glu Leu Gly Thr Leu Glu His Val His Ser	
35 40 45	
gga gca acg ccg cgg cga cgc ggc ctg acg agg cgc gag tca aac tcg	192
Gly Ala Thr Pro Arg Arg Arg Gly Leu Thr Arg Arg Glu Ser Asn Ser	
50 55 60	
gac gcg aac gac aac gat ccg ctg gtg gtc aac acg gat aag ggg cgc	240
Asp Ala Asn Asp Asn Asp Pro Leu Val Val Asn Thr Asp Lys Gly Arg	
65 70 75 80	
atc cgc ggc att acg gtc gat gcg ccc agc ggc aag aag gtg gac gtg	288
Ile Arg Gly Ile Thr Val Asp Ala Pro Ser Gly Lys Lys Val Asp Val	
85 90 95	
tgg ctc ggc att ccc tac gcc cag ccg ccg gtc ggg ccg cta cgg ttc	336
Trp Leu Gly Ile Pro Tyr Ala Gln Pro Pro Val Gly Pro Leu Arg Phe	
100 105 110	
cgt cat ccg cgg ccg gcc gaa aag tgg acc ggc gtg ctg aac acg acc	384
Arg His Pro Arg Pro Ala Glu Lys Trp Thr Gly Val Leu Asn Thr Thr	
115 120 125	
aca ccg ccc aac agc tgc gtg cag atc gtg gac acc gtg ttc ggc gac	432
Thr Pro Pro Asn Ser Cys Val Gln Ile Val Asp Thr Val Phe Gly Asp	
130 135 140	
ttc ccg ggc gcg acc atg tgg aac ccg aac acg ccc ctg tcc gag gac	480
Phe Pro Gly Ala Thr Met Trp Asn Pro Asn Thr Pro Leu Ser Glu Asp	
145 150 155 160	



tgt	ctg	tac	att	aac	gtg	gtg	gca	ccg	cga	ccc	cgg	ccc	aag	aat	gcg	528
Cys	Leu	Tyr	Ile	Asn	Val	Val	Ala	Pro	Arg	Pro	Arg	Pro	Lys	Asn	Ala	
				165					170						175	
gcc	gtc	atg	ctg	tgg	atc	ttc	ggc	ggc	ggc	ttc	tac	tcc	ggc	acc	gcc	576
Ala	Val	Met	Leu	Trp	Ile	Phe	Gly	Gly	Gly	Phe	Tyr	Ser	Gly	Thr	Ala	
			180					185					190			
acc	ctg	gac	gtg	tac	gac	cac	cgg	gcg	ctt	gcg	tcg	gag	gag	aac	gtg	624
Thr	Leu	Asp	Val	Tyr	Asp	His	Arg	Ala	Leu	Ala	Ser	Glu	Glu	Asn	Val	
		195					200					205				
atc	gtg	gtg	tcg	ctg	cag	tac	cgc	gtg	gcc	agt	ctg	ggc	ttc	ctg	ttt	672
Ile	Val	Val	Ser	Leu	Gln	Tyr	Arg	Val	Ala	Ser	Leu	Gly	Phe	Leu	Phe	
	210					215					220					
ctc	ggc	acc	ccg	gaa	gcg	ccg	ggc	aat	gcg	gga	ctg	ttc	gat	cag	aac	720
Leu	Gly	Thr	Pro	Glu	Ala	Pro	Gly	Asn	Ala	Gly	Leu	Phe	Asp	Gln	Asn	
225					230					235					240	
ctt	gcg	cta	cgc	tgg	gtg	cgg	gac	aac	att	cac	cgg	ttc	ggt	ggc	gat	768
Leu	Ala	Leu	Arg	Trp	Val	Arg	Asp	Asn	Ile	His	Arg	Phe	Gly	Gly	Asp	
				245					250					255		
ccg	tcg	cgt	gtg	aca	ctg	ttc	ggc	gag	agt	gcc	ggt	gcc	gtc	tcg	gtg	816
Pro	Ser	Arg	Val	Thr	Leu	Phe	Gly	Glu	Ser	Ala	Gly	Ala	Val	Ser	Val	
			260					265					270			
tcg	ctg	cat	ctg	ctg	tcc	gcc	ctt	tcc	cgc	gat	ctg	ttc	cag	cgg	gcc	864
Ser	Leu	His	Leu	Leu	Ser	Ala	Leu	Ser	Arg	Asp	Leu	Phe	Gln	Arg	Ala	
		275					280					285				
atc	ctg	cag	agc	ggc	tcg	ccg	acg	gca	ccg	tgg	gca	ttg	gta	tcg	cgc	912
Ile	Leu	Gln	Ser	Gly	Ser	Pro	Thr	Ala	Pro	Trp	Ala	Leu	Val	Ser	Arg	
	290					295				300						
gag	gaa	gcc	aca	cta	aga	gca	ctg	cgg	ttg	gcc	gag	gcg	gtc	ggc	tgc	960
Glu	Glu	Ala	Thr	Leu	Arg	Ala	Leu	Arg	Leu	Ala	Glu	Ala	Val	Gly	Cys	
305					310					315					320	
ccg	cac	gaa	ccg	agc	aag	ctg	agc	gat	gcg	gtc	gag	tgc	ctg	cgc	ggc	1008
Pro	His	Glu	Pro	Ser	Lys	Leu	Ser	Asp	Ala	Val	Glu	Cys	Leu	Arg	Gly	
				325				330						335		
aag	gac	ccg	cac	gtg	ctg	gtc	aac	aac	gag	tgg	ggc	acg	ctc	ggc	att	1056
Lys	Asp	Pro	His	Val	Leu	Val	Asn	Asn	Glu	Trp	Gly	Thr	Leu	Gly	Ile	
			340					345					350			
tgc	gag	ttc	ccg	ttc	gtg	ccg	gtg	gtc	gac	ggt	gcg	ttc	ctg	gac	gag	1104
Cys	Glu	Phe	Pro	Phe	Val	Pro	Val	Val	Asp	Gly	Ala	Phe	Leu	Asp	Glu	
		355					360					365				
acg	ccg	cag	cgt	tcg	ctc	gcc	agc	ggg	cgc	ttc	aag	aag	acg	gag	atc	1152
Thr	Pro	Gln	Arg	Ser	Leu	Ala	Ser	Gly	Arg	Phe	Lys	Lys	Thr	Glu	Ile	
	370					375					380					

ctc acc ggc agc aac acg gag gag ggc tac tac ttc atc atc tac tac	1200
Leu Thr Gly Ser Asn Thr Glu Glu Gly Tyr Tyr Phe Ile Ile Tyr Tyr	
385 390 395 400	
ctg acc gag ctg ctg cgc aag gag gag ggc gtg acc gtg acg cgc gag	1248
Leu Thr Glu Leu Arg Lys Glu Glu Gly Val Thr Val Thr Arg Glu	
405 410 415	
<del>gag ttc ctg cag gcg gtg cgc gag ctc aac ccg tac gtg aac ggg gcg</del>	<del>1296</del>
<del>Glu Phe Leu Gln Ala Val Arg Glu Leu Asn Pro Tyr Val Asn Gly Ala</del>	<del></del>
<del>420 425 430</del>	<del></del>
gcc cgg cag gcg atc gtg ttc gag tac acc gac tgg acc gag ccg gac	1344
Ala Arg Gln Ala Ile Val Phe Glu Tyr Thr Asp Trp Thr Glu Pro Asp	
435 440 445	
aac ccg aac agc aac cgg gac gcg ctg gac aag atg gtg ggc gac tat	1392
Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met Val Gly Asp Tyr	
450 455 460	
cac ttc acc tgc aac gtg aac gag ttc gcg cag cgg tac gcc gag gag	1440
His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg Tyr Ala Glu Glu	
465 470 475 480	
ggc aac aac gtc tac atg tat ctg tac acg cac cgc agc aaa ggc aac	1488
Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg Ser Lys Gly Asn	
485 490 495	
ccg tgg ccg cgc tgg acg ggc gtg atg cac ggc gac gag atc aac tac	1536
Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp Glu Ile Asn Tyr	
500 505 510	
gtg ttc ggc gaa ccg ctc aac ccc acc ctc ggc tac acc gag gac gag	1584
Val Phe Gly Glu Pro Leu Asn Pro Thr Leu Gly Tyr Thr Glu Asp Glu	
515 520 525	
aaa gac ttt agc cgg aag atc atg cga tac tgg tcc aac ttt gcc aaa	1632
Lys Asp Phe Ser Arg Lys Ile Met Arg Tyr Trp Ser Asn Phe Ala Lys	
530 535 540	
acc ggg aat cca aat ccc aac acg gcc agc agc gaa ttc ccc gag tgg	1680
Thr Gly Asn Pro Asn Pro Asn Thr Ala Ser Ser Glu Phe Pro Glu Trp	
545 550 555 560	
ccc aag cac acc gcc cac gga cgg cac tat ctg gag ctg ggc ctc aac	1728
Pro Lys His Thr Ala His Gly Arg His Tyr Leu Glu Leu Gly Leu Asn	
565 570 575	
acg tcc ttc gtc ggt cgg ggc cca cgg ttg agg cag tgt gcc ttc tgg	1776
Thr Ser Phe Val Gly Arg Gly Pro Arg Leu Arg Gln Cys Ala Phe Trp	
580 585 590	
aag aag tac ctt ccc cag cta gtt gca gct acc tcg aac cta cca ggg	1824
Lys Lys Tyr Leu Pro Gln Leu Val Ala Ala Thr Ser Asn Leu Pro Gly	
595 600 605	

cca gca ccg cct agt gaa ccg tgc gaa agc agc gca ttt ttt tac cga 1872  
 Pro Ala Pro Pro Ser Glu Pro Cys Glu Ser Ser Ala Phe Phe Tyr Arg  
 610 615 620

cct gat ctg atc gtg ctg ctg gtg tgc ctg ctt acg gcg acc gtc aga 1920  
 Pro Asp Leu Ile Val Leu Leu Val Ser Leu Leu Thr Ala Thr Val Arg  
 625 630 635 640

ttc ata caa taa 1932  
 Phe Ile Gln

<210> 3  
 <211> 643  
 <212> PRT  
 <213> Anopheles gambiae

<400> 3  
 Met Phe Val Cys Cys Phe Phe Phe Leu Ser Leu Ser Phe Cys Gly Ser  
 1 5 10 15  
 Asn Ile Ser Asp Ala Phe Phe Thr Pro Tyr Ile Gly His Gly Glu Ser  
 20 25 30  
 Val Arg Ile Ile Asp Ala Glu Leu Gly Thr Leu Glu His Val His Ser  
 35 40 45  
 Gly Ala Thr Pro Arg Arg Arg Gly Leu Thr Arg Arg Glu Ser Asn Ser  
 50 55 60  
 Asp Ala Asn Asp Asn Asp Pro Leu Val Val Asn Thr Asp Lys Gly Arg  
 65 70 75 80  
 Ile Arg Gly Ile Thr Val Asp Ala Pro Ser Gly Lys Lys Val Asp Val  
 85 90 95  
 Trp Leu Gly Ile Pro Tyr Ala Gln Pro Pro Val Gly Pro Leu Arg Phe  
 100 105 110  
 Arg His Pro Arg Pro Ala Glu Lys Trp Thr Gly Val Leu Asn Thr Thr  
 115 120 125  
 Thr Pro Pro Asn Ser Cys Val Gln Ile Val Asp Thr Val Phe Gly Asp  
 130 135 140  
 Phe Pro Gly Ala Thr Met Trp Asn Pro Asn Thr Pro Leu Ser Glu Asp  
 145 150 155 160  
 Cys Leu Tyr Ile Asn Val Val Ala Pro Arg Pro Arg Pro Lys Asn Ala  
 165 170 175  
 Ala Val Met Leu Trp Ile Phe Gly Gly Gly Phe Tyr Ser Gly Thr Ala  
 180 185 190  
 Thr Leu Asp Val Tyr Asp His Arg Ala Leu Ala Ser Glu Glu Asn Val  
 195 200 205

Ile Val Val Ser Leu Gln Tyr Arg Val Ala Ser Leu Gly Phe Leu Phe  
 210 215 220  
 Leu Gly Thr Pro Glu Ala Pro Gly Asn Ala Gly Leu Phe Asp Gln Asn  
 225 230 235 240  
 Leu Ala Leu Arg Trp Val Arg Asp Asn Ile His Arg Phe Gly Gly Asp  
 245 250 255  
 Pro Ser Arg Val Thr Leu Phe Gly Glu Ser Ala Gly Ala Val Ser Val  
 260 265 270  
 Ser Leu His Leu Leu Ser Ala Leu Ser Arg Asp Leu Phe Gln Arg Ala  
 275 280 285  
 Ile Leu Gln Ser Gly Ser Pro Thr Ala Pro Trp Ala Leu Val Ser Arg  
 290 295 300  
 Glu Glu Ala Thr Leu Arg Ala Leu Arg Leu Ala Glu Ala Val Gly Cys  
 305 310 315 320  
 Pro His Glu Pro Ser Lys Leu Ser Asp Ala Val Glu Cys Leu Arg Gly  
 325 330 335  
 Lys Asp Pro His Val Leu Val Asn Asn Glu Trp Gly Thr Leu Gly Ile  
 340 345 350  
 Cys Glu Phe Pro Phe Val Pro Val Val Asp Gly Ala Phe Leu Asp Glu  
 355 360 365  
 Thr Pro Gln Arg Ser Leu Ala Ser Gly Arg Phe Lys Lys Thr Glu Ile  
 370 375 380  
 Leu Thr Gly Ser Asn Thr Glu Glu Gly Tyr Tyr Phe Ile Ile Tyr Tyr  
 385 390 395 400  
 Leu Thr Glu Leu Leu Arg Lys Glu Glu Gly Val Thr Val Thr Arg Glu  
 405 410 415  
 Glu Phe Leu Gln Ala Val Arg Glu Leu Asn Pro Tyr Val Asn Gly Ala  
 420 425 430  
 Ala Arg Gln Ala Ile Val Phe Glu Tyr Thr Asp Trp Thr Glu Pro Asp  
 435 440 445  
 Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met Val Gly Asp Tyr  
 450 455 460  
 His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg Tyr Ala Glu Glu  
 465 470 475 480  
 Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg Ser Lys Gly Asn  
 485 490 495  
 Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp Glu Ile Asn Tyr  
 500 505 510

Val Phe Gly Glu Pro Leu Asn Pro Thr Leu Gly Tyr Thr Glu Asp Glu  
515 520 525

Lys Asp Phe Ser Arg Lys Ile Met Arg Tyr Trp Ser Asn Phe Ala Lys  
530 535 540

Thr Gly Asn Pro Asn Pro Asn Thr Ala Ser Ser Glu Phe Pro Glu Trp  
545 550 555 560

Pro Lys His Thr Ala His Gly Arg His Tyr Leu Glu Leu Gly Leu Asn  
565 570 575

Thr Ser Phe Val Gly Arg Gly Pro Arg Leu Arg Gln Cys Ala Phe Trp  
580 585 590

Lys Lys Tyr Leu Pro Gln Leu Val Ala Ala Thr Ser Asn Leu Pro Gly  
595 600 605

Pro Ala Pro Pro Ser Glu Pro Cys Glu Ser Ser Ala Phe Phe Tyr Arg  
610 615 620

Pro Asp Leu Ile Val Leu Leu Val Ser Leu Leu Thr Ala Thr Val Arg  
625 630 635 640

Phe Ile Gln

<210> 4

<211> 1932

<212> ADN

<213> Anopheles gambiae souche KISUMU

<220>

<221> CDS

<222> (1)..(1932)

<400> 4

atg ttt gtg tgt tgt ttt ttc ttt ctc tct ctc tct ctc tgt ggt tcc 48  
Met Phe Val Cys Cys Phe Phe Phe Leu Ser Leu Ser Leu Cys Gly Ser  
1 5 10 15

aac att tca gac gca ttt ttt aca cca tat ata ggt cac ggt gag tcc 96  
Asn Ile Ser Asp Ala Phe Phe Thr Pro Tyr Ile Gly His Gly Glu Ser  
20 25 30

gta cga att ata gat gcc gag ttg ggc acg ctc gag cat gtc cac agt 144  
Val Arg Ile Ile Asp Ala Glu Leu Gly Thr Leu Glu His Val His Ser  
35 40 45

gga gca acg ccg cgg cga cgc ggt ctg acg agg cgc gag tcc aac tcg 192  
Gly Ala Thr Pro Arg Arg Arg Gly Leu Thr Arg Arg Glu Ser Asn Ser  
50 55 60

gac gcg aac gac aac gat ccg ctg gtg gtc aac acg gat aag ggg cgc 240  
Asp Ala Asn Asp Asn Asp Pro Leu Val Val Asn Thr Asp Lys Gly Arg  
65 70 75 80

atc cgc ggc att acg gtc gat gcg ccc agc ggc aag aag gtg gac gtg	288
Ile Arg Gly Ile Thr Val Asp Ala Pro Ser Gly Lys Lys Val Asp Val	
85 90 95	
tgg ctc ggc att ccc tac gcc cag ccg ccg gtc ggg ccg tta cgg ttc	336
Trp Leu Gly Ile Pro Tyr Ala Gln Pro Pro Val Gly Pro Leu Arg Phe	
100 105 110	
cgt cat ccg cgg ccg gcc gaa aag tgg acc ggc gtg ctg aac acg acc	384
Arg His Pro Arg Pro Ala Glu Lys Trp Thr Gly Val Leu Asn Thr Thr	
115 120 125	
aca ccg ccc aac agc tgc gtg cag atc gtg gac acc gtg ttc ggc gac	432
Thr Pro Pro Asn Ser Cys Val Gln Ile Val Asp Thr Val Phe Gly Asp	
130 135 140	
ttc ccg ggc gcg acc atg tgg aac ccg aac acg ccc ctg tcc gag gac	480
Phe Pro Gly Ala Thr Met Trp Asn Pro Asn Thr Pro Leu Ser Glu Asp	
145 150 155 160	
tgt ctg tac att aac gtg gtg gca ccg cga ccc cgg ccc aag aat gcg	528
Cys Leu Tyr Ile Asn Val Val Ala Pro Arg Pro Arg Pro Lys Asn Ala	
165 170 175	
gcc gtc atg ctg tgg atc ttc ggc ggc ggc ttc tac tcc ggc acc gcc	576
Ala Val Met Leu Trp Ile Phe Gly Gly Gly Phe Tyr Ser Gly Thr Ala	
180 185 190	
acc ctg gac gtg tac gac cac ccg gcg ctt gcg tcg gag gag aac gtg	624
Thr Leu Asp Val Tyr Asp His Arg Ala Leu Ala Ser Glu Glu Asn Val	
195 200 205	
atc gtg gtg tcg ctg cag tac cgc gtg gcc agt ctg ggc ttc ctg ttt	672
Ile Val Val Ser Leu Gln Tyr Arg Val Ala Ser Leu Gly Phe Leu Phe	
210 215 220	
ctc ggc acc ccg gaa gcg ccg ggc aat gcg gga ctg ttc gat cag aac	720
Leu Gly Thr Pro Glu Ala Pro Gly Asn Ala Gly Leu Phe Asp Gln Asn	
225 230 235 240	
ctt gcg cta cgc tgg gtg ccg gac aac att cac ccg ttc ggt ggt gat	768
Leu Ala Leu Arg Trp Val Arg Asp Asn Ile His Arg Phe Gly Gly Asp	
245 250 255	
ccg tcg cgt gtg aca ctg ttc ggc gag agt gcc ggt gcc gtc tcg gtg	816
Pro Ser Arg Val Thr Leu Phe Gly Glu Ser Ala Gly Ala Val Ser Val	
260 265 270	
tcg ctg cat ctg ctg tcc gcc ctg tcc cgc gat ctg ttc cag ccg gcc	864
Ser Leu His Leu Leu Ser Ala Leu Ser Arg Asp Leu Phe Gln Arg Ala	
275 280 285	
atc ctg cag agc ggc tcg ccg acg gca ccg tgg gca ttg gta tcg cgc	912
Ile Leu Gln Ser Gly Ser Pro Thr Ala Pro Trp Ala Leu Val Ser Arg	
290 295 300	
gag gaa gcc acg cta aga gca ctg ccg ttg gcc gag gcg gtc ggc tgc	960
Glu Glu Ala Thr Leu Arg Ala Leu Arg Leu Ala Glu Ala Val Gly Cys	

305	310	315	320	
ccg cac gaa ccg agc aag ctg agc gat gcg gtc gag tgt ctg cgc ggc				1008
Pro His Glu Pro Ser Lys Leu Ser Asp Ala Val Glu Cys Leu Arg Gly				
	325	330	335	
aag gat ccg cac gtg ctg gtc aac aac gag tgg ggc acg ctc ggc att				1056
Lys Asp Pro His Val Leu Val Asn Asn Glu Trp Gly Thr Leu Gly Ile				
	340	345	350	
tgc gag ttc ccg ttc gtg ccg gtg gtc gac ggt gcg ttc ctg gac gag				1104
Cys Glu Phe Pro Phe Val Pro Val Val Asp Gly Ala Phe Leu Asp Glu				
	355	360	365	
acg ccg cag cgt tgc ctc gcc agc ggg cgc ttc aag aag acg gag atc				1152
Thr Pro Gln Arg Ser Leu Ala Ser Gly Arg Phe Lys Lys Thr Glu Ile				
	370	375	380	
ctc acc ggc agc aac acg gag gag ggc tac tac ttc atc atc tac tac				1200
Leu Thr Gly Ser Asn Thr Glu Glu Gly Tyr Tyr Phe Ile Ile Tyr Tyr				
	385	390	395	400
ctg acc gag ctg ctg cgc aag gag gag ggc gtg acc gtg acg cgc gag				1248
Leu Thr Glu Leu Leu Arg Lys Glu Glu Gly Val Thr Val Thr Arg Glu				
	405	410	415	
gag ttc ctg cag gcg gtg cgc gag ctc aac ccg tac gtg aac ggg gcg				1296
Glu Phe Leu Gln Ala Val Arg Glu Leu Asn Pro Tyr Val Asn Gly Ala				
	420	425	430	
gcc cgg cag gcg atc gtg ttc gag tac acc gac tgg acc gag ccg gac				1344
Ala Arg Gln Ala Ile Val Phe Glu Tyr Thr Asp Trp Thr Glu Pro Asp				
	435	440	445	
aac ccg aac agc aac ccg gac gcg ctg gac aag atg gtg ggc gac tat				1392
Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met Val Gly Asp Tyr				
	450	455	460	
cac ttc acc tgc aac gtg aac gag ttc gcg cag cgg tac gcc gag gag				1440
His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg Tyr Ala Glu Glu				
	465	470	475	480
ggc aac aac gtc tac atg tat ctg tac acg cac cgc agc aaa ggc aac				1488
Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg Ser Lys Gly Asn				
	485	490	495	
ccg tgg ccg cgc tgg acg ggc gtg atg cac ggc gac gag atc aac tac				1536
Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp Glu Ile Asn Tyr				
	500	505	510	
gtg ttc ggc gaa ccg ctc aac ccc acc ctc ggc tac acc gag gac gag				1584
Val Phe Gly Glu Pro Leu Asn Pro Thr Leu Gly Tyr Thr Glu Asp Glu				
	515	520	525	
aaa gac ttt agc ccg aag atc atg cga tac tgg tct aac ttt gcc aaa				1632
Lys Asp Phe Ser Arg Lys Ile Met Arg Tyr Trp Ser Asn Phe Ala Lys				
	530	535	540	

acc ggg aat cca aat ccc aac acg gcc agc agc gaa ttc ccc gag tgg 1680  
 Thr Gly Asn Pro Asn Pro Asn Thr Ala Ser Ser Glu Phe Pro Glu Trp  
 545 550 555 560

ccc aag cac acc gcc cac gga cgg cac tat ctg gag ctg ggc ctc aac 1728  
 Pro Lys His Thr Ala His Gly Arg His Tyr Leu Glu Leu Gly Leu Asn  
 565 570 575

acg tcc ttc gtc ggt cgg ggc cca cgg ttg agg cag tgt gcc ttc tgg 1776  
 Thr Ser Phe Val Gly Arg Gly Pro Arg Leu Arg Gln Cys Ala Phe Trp  
 580 585 590

aag aag tac ctt ccc cag cta gtt gca gct acc tcg aac cta cca ggg 1824  
 Lys Lys Tyr Leu Pro Gln Leu Val Ala Ala Thr Ser Asn Leu Pro Gly  
 595 600 605

cca gca ccg ccc agt gaa ccg tgc gaa agc agc gca ttt ttt tac cga 1872  
 Pro Ala Pro Pro Ser Glu Pro Cys Glu Ser Ser Ala Phe Phe Tyr Arg  
 610 615 620

cct gat ctg atc gtg ctg ctg gtg tgc ctg ctt acg gcg acc gtc aga 1920  
 Pro Asp Leu Ile Val Leu Leu Val Ser Leu Leu Thr Ala Thr Val Arg  
 625 630 635 640

ttc ata caa taa 1932  
 Phe Ile Gln

<210> 5  
 <211> 643  
 <212> PRT  
 <213> Anopheles gambiae souche KISUMU

<400> 5  
 Met Phe Val Cys Cys Phe Phe Phe Leu Ser Leu Ser Leu Cys Gly Ser  
 1 5 10 15

Asn Ile Ser Asp Ala Phe Phe Thr Pro Tyr Ile Gly His Gly Glu Ser  
 20 25 30

Val Arg Ile Ile Asp Ala Glu Leu Gly Thr Leu Glu His Val His Ser  
 35 40 45

Gly Ala Thr Pro Arg Arg Arg Gly Leu Thr Arg Arg Glu Ser Asn Ser  
 50 55 60

Asp Ala Asn Asp Asn Asp Pro Leu Val Val Asn Thr Asp Lys Gly Arg  
 65 70 75 80

Ile Arg Gly Ile Thr Val Asp Ala Pro Ser Gly Lys Lys Val Asp Val  
 85 90 95

Trp Leu Gly Ile Pro Tyr Ala Gln Pro Pro Val Gly Pro Leu Arg Phe  
 100 105 110

Arg His Pro Arg Pro Ala Glu Lys Trp Thr Gly Val Leu Asn Thr Thr  
 115 120 125



Thr Pro Pro Asn Ser Cys Val Gln Ile Val Asp Thr Val Phe Gly Asp  
 130 135 140  
 Phe Pro Gly Ala Thr Met Trp Asn Pro Asn Thr Pro Leu Ser Glu Asp  
 145 150 155 160  
 Cys Leu Tyr Ile Asn Val Val Ala Pro Arg Pro Arg Pro Lys Asn Ala  
 165 170 175  
 Ala Val Met Leu Trp Ile Phe Gly Gly Gly Phe Tyr Ser Gly Thr Ala  
 180 185 190  
 Thr Leu Asp Val Tyr Asp His Arg Ala Leu Ala Ser Glu Glu Asn Val  
 195 200 205  
 Ile Val Val Ser Leu Gln Tyr Arg Val Ala Ser Leu Gly Phe Leu Phe  
 210 215 220  
 Leu Gly Thr Pro Glu Ala Pro Gly Asn Ala Gly Leu Phe Asp Gln Asn  
 225 230 235 240  
 Leu Ala Leu Arg Trp Val Arg Asp Asn Ile His Arg Phe Gly Gly Asp  
 245 250 255  
 Pro Ser Arg Val Thr Leu Phe Gly Glu Ser Ala Gly Ala Val Ser Val  
 260 265 270  
 Ser Leu His Leu Leu Ser Ala Leu Ser Arg Asp Leu Phe Gln Arg Ala  
 275 280 285  
 Ile Leu Gln Ser Gly Ser Pro Thr Ala Pro Trp Ala Leu Val Ser Arg  
 290 295 300  
 Glu Glu Ala Thr Leu Arg Ala Leu Arg Leu Ala Glu Ala Val Gly Cys  
 305 310 315 320  
 Pro His Glu Pro Ser Lys Leu Ser Asp Ala Val Glu Cys Leu Arg Gly  
 325 330 335  
 Lys Asp Pro His Val Leu Val Asn Asn Glu Trp Gly Thr Leu Gly Ile  
 340 345 350  
 Cys Glu Phe Pro Phe Val Pro Val Val Asp Gly Ala Phe Leu Asp Glu  
 355 360 365  
 Thr Pro Gln Arg Ser Leu Ala Ser Gly Arg Phe Lys Lys Thr Glu Ile  
 370 375 380  
 Leu Thr Gly Ser Asn Thr Glu Glu Gly Tyr Tyr Phe Ile Ile Tyr Tyr  
 385 390 395 400  
 Leu Thr Glu Leu Leu Arg Lys Glu Glu Gly Val Thr Val Thr Arg Glu  
 405 410 415  
 Glu Phe Leu Gln Ala Val Arg Glu Leu Asn Pro Tyr Val Asn Gly Ala  
 420 425 430

Ala Arg Gln Ala Ile Val Phe Glu Tyr Thr Asp Trp Thr Glu Pro Asp  
 435 440 445

Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met Val Gly Asp Tyr  
 450 455 460

His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg Tyr Ala Glu Glu  
 465 470 475 480

~~Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg Ser Lys Gly Asn~~  
 485 490 495

Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp Glu Ile Asn Tyr  
 500 505 510

Val Phe Gly Glu Pro Leu Asn Pro Thr Leu Gly Tyr Thr Glu Asp Glu  
 515 520 525

Lys Asp Phe Ser Arg Lys Ile Met Arg Tyr Trp Ser Asn Phe Ala Lys  
 530 535 540

Thr Gly Asn Pro Asn Pro Asn Thr Ala Ser Ser Glu Phe Pro Glu Trp  
 545 550 555 560

Pro Lys His Thr Ala His Gly Arg His Tyr Leu Glu Leu Gly Leu Asn  
 565 570 575

Thr Ser Phe Val Gly Arg Gly Pro Arg Leu Arg Gln Cys Ala Phe Trp  
 580 585 590

Lys Lys Tyr Leu Pro Gln Leu Val Ala Ala Thr Ser Asn Leu Pro Gly  
 595 600 605

Pro Ala Pro Pro Ser Glu Pro Cys Glu Ser Ser Ala Phe Phe Tyr Arg  
 610 615 620

Pro Asp Leu Ile Val Leu Leu Val Ser Leu Leu Thr Ala Thr Val Arg  
 625 630 635 640

Phe Ile Gln

<210> 6

<211> 3297

<212> ADN

<213> Culex pipiens souche S-LAB

<400> 6

ccagagcaga ccacgaacct cgtcggaaga gctgatgccg ttgtgacatt cgctccgatt 60  
 gtgtaagcaa ataaggtag gacacaccgt attcacgaac tctgacacca agctgtcata 120  
 gccgtcactg acgagaagaa aaagaaacaa gagtcgacaa cacactcaca gtctcacgcc 180  
 gccagagagc acaccaagag tcacattgag aaaaccacac gccagaagaa aagaagagtt 240  
 gttcaagaag gaagctaata ccacacacac acacactcac acacaccggg agaaaccgca 300  
 cagcaggcgg cgctgtgaaa ttcacacggt cggtcgggtga agtggtggaa ggaactcggc 360  
 gtcggagtag caattagtag attacaaaca aagggaaata agggaggag tcaagagtca 420  
 accagtggaa ccagtggtag agtgagtgat ttttttgtgt tggtgctgca gaaaggaacg 480  
 cgcgacgagc acactcttgt gaaatcgggt tcatcatcgt taaatgctct cgaccgtcaa 540

```

cttatagcta tcatatgcga tctctccaag ccatggagat ccgaggccta ataaccgat 600
tactgggtcc atgtcacctg cgacatctga tactgtgcag tttggggctg tactccatcc 660
tcgtgaagtc ggtccattgc cggcatcatg acatcggtag ttcgggtggca caccagctag 720
gatcgaaata ctcacaatca tcctcgttat cgtcatcctc gcaatcgta tcgtcgtag 780
ctgaagaggc cacgctgaat aaagattcag atgcattttt tacaccatat ataggtcacg 840
gagattctgt tcgaattgta gatgccgaat taggtacatt agagcgcgag cacatccata 900
gcactacgac ccggcggcgt ggcctgacgc ggaggggagtc cagctccgat gccaccgact 960
cggaccact ggtcataacg acggacaagg gcaaaatccg tggaacgaca ctggaagcgc 1020
ctagtggaaa gaaggtggac gcatggatgg gcattccgta cgcgcagccc ccgctgggtc 1080
cgctccggtt tcgacatccg cgaccggccg aaagatggac cgggtgtgctg aaccgcacca 1140
aaccgccccaa ctctgcgtc cagatcgtgg acaccgtgtt cggtgacttc ccgggggcca 1200
ccatgtggaa cccgaacaca ccgctctcgg aggactgtct gtacatcaac gtggtcgtgc 1260
cacggcccag gcccaagaat gccgccgtca tgctgtggat cttcgggggt ggcttctact 1320
ccgggactgc cacgctggac gtgtacgacc atcggacgct ggcctcggag gagaacgtga 1380
tcgtagtttc gctgcagtac cgtgtcgcaa gtcttgggtt tctcttcctc ggcaaccgg 1440
aggcaccgg taacgcgggg ctgtttgatc agaacctggc actgagatgg gtccgcgaca 1500
acatccaccg gttcggcggg gacccctcgc gggtcacact gttcggcgag agcgcggag 1560
cggctctcggg ttcgctgcac ctgctgtcgg cgctctcggg ggacctgttc cagcgggcca 1620
tcctccagag tggctccccg acggcccgt gggcgctggg ttcgcgcgaa gaagctacgc 1680
ttagagctct tcgtctggcc gagccgtca actgtccgca cgtatcgacc aagctgagcg 1740
atgccgtcga atgcctgcga accaaggatc cgaacgagct ggtcgacaac gagggggca 1800
cgctggggat ctgcgagttt ccgttcgttc cggttgtgga cggagccttc ctcgatgaga 1860
caccgcagcg ttcgttggcc agcgggcgct tcaagaaaac ggacatcctg accggcagca 1920
acaccgagga gggttactac tttatcattt actatctaac cgagctgctc aggaaagagg 1980
aaggggtcac ggtaacacgc gaggagttcc tacaggccgt ccgggagttg aatccgtacg 2040
tgaacgggtg cgcccggcag gccatcgtgt tcgagtacac ggactggatt gaaccggaca 2100
acccgaacag caaccgtgac gcgctggaca agatggtcgg ggattatcac ttcacctgca 2160
acgtgaacga attcgcccag cggtagccg aggagggcaa caacgtgttc atgtacctgt 2220
acacgcacag aagcaaaagg aatccctggc cgaggtggac cggcgtgatg cagggcgacg 2280
agatcaacta cgtgtttggc gaaccgtga actcggccct cggctaccag gacgacgaga 2340
aggactttag ccgaaaaatt atgcgatact ggtccaactt tgccaagact ggcaatccca 2400
acccgagtac gccgagcgtg gacctgcccg aatggcccaa gcacaccgcc cacggacgac 2460
actatctgga gctgggactg aacacgacct tcgtgggacg gggccacga ttgcggcagt 2520
gcgctttctg gaagaaatat ttgccgaac tagtagcagc tacctctaac ctccaagtaa 2580
ctccgcgcc tagcgtacct tgcgaaagca gctcaacatc ttatcgatcc actctacttc 2640
taatagtcac actactttta gtaacgcggt tcaagattta aatccgtgtt ttctttcccg 2700
ttcccgtttt tcggttaaag cttcttttagg tcaggtgaaa acatcaacaa gcagcatcaa 2760
ttctactact aatactatta ctactattaa ctgaaatgga acaataagat tacttttttc 2820
ttctaaattt gttcaactgc taattaaatt ctaaataggt gaatgcatct tgctctgcaa 2880
acgaacgata ggacaattat gttgtattgt ttttttcttt gtaataatat tctgtaaaaa 2940
gaggtgatata cattaatatt ttactaacca tacaataaac aaaatatttc ctgttataaa 3000
ttgtgatgaa tatttcgctt taactacacc attgaagggt acttaagttg aaataacaaa 3060
aattttatat aaacaactaa caaataaaaac agctgctaga gacaactaga cattaatatc 3120
aaaaaaacgt tattttgaaa aagagcgatt tatgcactag cggagggtgaa tcccttataa 3180
tcttgaaaag agaggaggaa tggaagaaga agaagaagaa aatattatga tacaataaaa 3240
ccaacatcta attctaacaa tcaactgttt actttactaa aaaaaaaaaa aaaaaaa 3297

```

&lt;210&gt; 7

&lt;211&gt; 702

&lt;212&gt; PRT

&lt;213&gt; Culex pipiens souche S-LAB

&lt;400&gt; 7

Met Glu Ile Arg Gly Leu Ile Thr Arg Leu Leu Gly Pro Cys His Leu

1

5

10

15

Arg His Leu Ile Leu Cys Ser Leu Gly Leu Tyr Ser Ile Leu Val Lys  
                     20                    25                    30  
 Ser Val His Cys Arg His His Asp Ile Gly Ser Ser Val Ala His Gln  
                     35                    40                    45  
 Leu Gly Ser Lys Tyr Ser Gln Ser Ser Ser Leu Ser Ser Ser Ser Gln  
                     50                    55                    60  


---

 Ser Ser Ser Ser Leu Ala Glu Glu Ala Thr Leu Asn Lys Asp Ser Asp  
                     65                    70                    75                    80  
 Ala Phe Phe Thr Pro Tyr Ile Gly His Gly Asp Ser Val Arg Ile Val  
                     85                    90                    95  
 Asp Ala Glu Leu Gly Thr Leu Glu Arg Glu His Ile His Ser Thr Thr  
                     100                    105                    110  
 Thr Arg Arg Arg Gly Leu Thr Arg Arg Glu Ser Ser Ser Asp Ala Thr  
                     115                    120                    125  
 Asp Ser Asp Pro Leu Val Ile Thr Thr Asp Lys Gly Lys Ile Arg Gly  
                     130                    135                    140  
 Thr Thr Leu Glu Ala Pro Ser Gly Lys Lys Val Asp Ala Trp Met Gly  
                     145                    150                    155                    160  
 Ile Pro Tyr Ala Gln Pro Pro Leu Gly Pro Leu Arg Phe Arg His Pro  
                     165                    170                    175  
 Arg Pro Ala Glu Arg Trp Thr Gly Val Leu Asn Ala Thr Lys Pro Pro  
                     180                    185                    190  
 Asn Ser Cys Val Gln Ile Val Asp Thr Val Phe Gly Asp Phe Pro Gly  
                     195                    200                    205  
 Ala Thr Met Trp Asn Pro Asn Thr Pro Leu Ser Glu Asp Cys Leu Tyr  
                     210                    215                    220  
 Ile Asn Val Val Val Pro Arg Pro Arg Pro Lys Asn Ala Ala Val Met  
                     225                    230                    235                    240  
 Leu Trp Ile Phe Gly Gly Gly Phe Tyr Ser Gly Thr Ala Thr Leu Asp  
                     245                    250                    255  
 Val Tyr Asp His Arg Thr Leu Ala Ser Glu Glu Asn Val Ile Val Val  
                     260                    265                    270  
 Ser Leu Gln Tyr Arg Val Ala Ser Leu Gly Phe Leu Phe Leu Gly Thr  
                     275                    280                    285  
 Pro Glu Ala Pro Gly Asn Ala Gly Leu Phe Asp Gln Asn Leu Ala Leu  
                     290                    295                    300  
 Arg Trp Val Arg Asp Asn Ile His Arg Phe Gly Gly Asp Pro Ser Arg  
                     305                    310                    315                    320

Val Thr Leu Phe Gly Glu Ser Ala Gly Ala Val Ser Val Ser Leu His  
 325 330 335  
 Leu Leu Ser Ala Leu Ser Arg Asp Leu Phe Gln Arg Ala Ile Leu Gln  
 340 345 350  
 Ser Gly Ser Pro Thr Ala Pro Trp Ala Leu Val Ser Arg Glu Glu Ala  
 355 360 365  
 Thr Leu Arg Ala Leu Arg Leu Ala Glu Ala Val Asn Cys Pro His Asp  
 370 375 380  
 Ala Thr Lys Leu Ser Asp Ala Val Glu Cys Leu Arg Thr Lys Asp Pro  
 385 390 395 400  
 Asn Glu Leu Val Asp Asn Glu Trp Gly Thr Leu Gly Ile Cys Glu Phe  
 405 410 415  
 Pro Phe Val Pro Val Val Asp Gly Ala Phe Leu Asp Glu Thr Pro Gln  
 420 425 430  
 Arg Ser Leu Ala Ser Gly Arg Phe Lys Lys Thr Asp Ile Leu Thr Gly  
 435 440 445  
 Ser Asn Thr Glu Glu Gly Tyr Tyr Phe Ile Ile Tyr Tyr Leu Thr Glu  
 450 455 460  
 Leu Leu Arg Lys Glu Glu Gly Val Thr Val Thr Arg Glu Glu Phe Leu  
 465 470 475 480  
 Gln Ala Val Arg Glu Leu Asn Pro Tyr Val Asn Gly Ala Ala Arg Gln  
 485 490 495  
 Ala Ile Val Phe Glu Tyr Thr Asp Trp Ile Glu Pro Asp Asn Pro Asn  
 500 505 510  
 Ser Asn Arg Asp Ala Leu Asp Lys Met Val Gly Asp Tyr His Phe Thr  
 515 520 525  
 Cys Asn Val Asn Glu Phe Ala Gln Arg Tyr Ala Glu Glu Gly Asn Asn  
 530 535 540  
 Val Phe Met Tyr Leu Tyr Thr His Arg Ser Lys Gly Asn Pro Trp Pro  
 545 550 555 560  
 Arg Trp Thr Gly Val Met His Gly Asp Glu Ile Asn Tyr Val Phe Gly  
 565 570 575  
 Glu Pro Leu Asn Ser Ala Leu Gly Tyr Gln Asp Asp Glu Lys Asp Phe  
 580 585 590  
 Ser Arg Lys Ile Met Arg Tyr Trp Ser Asn Phe Ala Lys Thr Gly Asn  
 595 600 605  
 Pro Asn Pro Ser Thr Pro Ser Val Asp Leu Pro Glu Trp Pro Lys His  
 610 615 620

Thr Ala His Gly Arg His Tyr Leu Glu Leu Gly Leu Asn Thr Thr Phe  
625 630 635 640

Val Gly Arg Gly Pro Arg Leu Arg Gln Cys Ala Phe Trp Lys Lys Tyr  
645 650 655

Leu Pro Gln Leu Val Ala Ala Thr Ser Asn Leu Gln Val Thr Pro Ala  
660 665 670

~~Pro Ser Val Pro Cys Glu Ser Ser Ser Thr Ser Tyr Arg Ser Thr Leu~~  
675 680 685

Leu Leu Ile Val Thr Leu Leu Leu Val Thr Arg Phe Lys Ile  
690 695 700

<210> 8  
<211> 91  
<212> PRT  
<213> Culex pipiens

<400> 8  
Ile Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
1 5 10 15  
Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
20 25 30  
Tyr Ala Glu Glu Gly Asn Asn Val Phe Met Tyr Leu Tyr Thr His Arg  
35 40 45  
Ser Lys Gly Asn Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp  
50 55 60  
Glu Ile Asn Tyr Val Phe Gly Glu Pro Leu Asn Ser Ala Leu Gly Tyr  
65 70 75 80  
Gln Asp Asp Glu Lys Asp Phe Ser Arg Lys Ile  
85 90

<210> 9  
<211> 91  
<212> PRT  
<213> Aedes aegypti

<400> 9  
Thr Glu Pro Glu Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
1 5 10 15  
Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
20 25 30  
Tyr Ala Glu Glu Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg  
35 40 45

1

•

• •

Thr Asp Asp Glu Lys Gly Phe Ser Arg Lys Ile  
                             85                            90

<210> 12  
 <211> 91  
 <212> PRT  
 <213> Anopheles sundaicus

<400> 12  
 Thr Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
   1                            5                            10                            15  
 Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
                             20                            25                            30  
 Tyr Ala Glu Glu Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg  
                             35                            40                            45  
 Ser Lys Gly Asn Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp  
                             50                            55                            60  
 Glu Ile Asn Tyr Val Phe Gly Glu Pro Leu Asn Pro Thr Leu Gly Tyr  
   65                            70                            75                            80  
 Thr Glu Asp Glu Lys Asp Phe Ser Arg Lys Ile  
                             85                            90

<210> 13  
 <211> 91  
 <212> PRT  
 <213> Anopheles minimus

<400> 13  
 Thr Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
   1                            5                            10                            15  
 Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
                             20                            25                            30  
 Tyr Ala Glu Glu Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg  
                             35                            40                            45  
 Ser Lys Gly Asn Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp  
                             50                            55                            60  
 Glu Ile Asn Tyr Val Phe Gly Glu Pro Leu Asn Pro Ser Leu Gly Tyr  
   65                            70                            75                            80  
 Thr Glu Asp Glu Lys Asp Phe Ser Arg Lys Ile  
                             85                            90

<210> 14



<211> 91  
 <212> PRT  
 <213> Anopheles moucheti

<400> 14  
 Thr Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
   1                  5                  10                  15  
 Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
                   20                  25                  30  
 Tyr Ala Glu Glu Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg  
                   35                  40                  45  
 Ser Lys Gly Asn Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp  
                   50                  55                  60  
 Glu Ile Asn Tyr Val Phe Gly Glu Pro Leu Asn Pro Ser Leu Gly Tyr  
   65                  70                  75                  80  
 Thr Glu Asp Glu Lys Asp Phe Ser Arg Lys Ile  
                   85                  90

<210> 15  
 <211> 91  
 <212> PRT  
 <213> Anopheles arabiensis

<400> 15  
 Thr Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
   1                  5                  10                  15  
 Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
                   20                  25                  30  
 Tyr Ala Glu Glu Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg  
                   35                  40                  45  
 Ser Lys Gly Asn Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp  
                   50                  55                  60  
 Glu Ile Asn Tyr Val Phe Gly Glu Pro Leu Asn Pro Thr Leu Gly Tyr  
   65                  70                  75                  80  
 Thr Glu Asp Glu Lys Asp Phe Ser Arg Lys Ile  
                   85                  90

<210> 16  
 <211> 91  
 <212> PRT  
 <213> Anopheles funestus

&lt;400&gt; 16

Thr Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
 1 5 10 15

Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
 20 25 30

Tyr Ala Glu Glu Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg  
 35 40 45

---

Ser Lys Gly Asn Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp  
 50 55 60

Glu Ile Asn Tyr Val Phe Gly Glu Pro Leu Asn Pro Ser Leu Gly Tyr  
 65 70 75 80

Thr Glu Asp Glu Lys Asp Phe Ser Arg Lys Ile  
 85 90

&lt;210&gt; 17

&lt;211&gt; 91

&lt;212&gt; PRT

&lt;213&gt; Anopheles pseudopunctipennis

&lt;400&gt; 17

Thr Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
 1 5 10 15

Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
 20 25 30

Tyr Ala Glu Glu Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg  
 35 40 45

Ser Lys Gly Asn Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp  
 50 55 60

Glu Ile Asn Tyr Val Phe Gly Glu Pro Leu Asn Pro Gly Leu Gly Tyr  
 65 70 75 80

Thr Glu Asp Glu Lys Asp Phe Ser Arg Lys Ile  
 85 90

&lt;210&gt; 18

&lt;211&gt; 91

&lt;212&gt; PRT

&lt;213&gt; Anopheles sacharovi

&lt;400&gt; 18

Thr Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
 1 5 10 15

Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
 20 25 30

!

1

```

<400> 20
Thr Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met
  1                    5                10              15

Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg
      20                25              30

Tyr Ala Glu Glu Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg
      35                40              45

Ser Lys Gly Asn Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp
  50                    55                60

```

Glu Ile Asn Tyr Val Phe Gly Glu Pro Leu Asn Pro Thr Leu Gly Tyr  
 65 70 75 80

Thr Asp Asp Glu Lys Gly Phe Ser Arg Lys Ile  
 85 90

<210> 21

<211> 91

<212> PRT

<213> Anopheles nili

<400> 21

Thr Glu Pro Asp Asn Pro Asn Ser Asn Arg Asp Ala Leu Asp Lys Met  
 1 5 10 15

Val Gly Asp Tyr His Phe Thr Cys Asn Val Asn Glu Phe Ala Gln Arg  
 20 25 30

Tyr Ala Glu Gly Asn Asn Val Tyr Met Tyr Leu Tyr Thr His Arg  
 35 40 45

Ser Lys Gly Asn Pro Trp Pro Arg Trp Thr Gly Val Met His Gly Asp  
 50 55 60

Glu Ile Asn Tyr Val Phe Gly Glu Pro Leu Asn Pro Ser Leu Gly Tyr  
 65 70 75 80

Thr Glu Asp Glu Lys Asp Phe Ser Arg Lys Met  
 85 90

<210> 22

<211> 4209

<212> ADN

<213> Anopheles gambiae

<400> 22

tggtaattac aattcccaag tttgcgtatg acaatgttaa atgttaagac gctcaaatgc 60  
 aaccaataga gtataattac taaggcgggc agtagaaacc aaaatatctt aaataatgtc 120  
 aagcaaaaaca aaaagaacaa ttccggtcac tgctcaaaga aagccctaac taactaccta 180  
 accttttcat cgatgaccct gtactgacat ggtaagatat tctttatcct ttaactcttc 240  
 tgcaccctac gcaactcaatg caacacacgc actactatta ctgctactac tctcgcactc 300  
 acgagcacct acttgcactc aagccggcac tcaatgtact agcgaacac gtcgcatcta 360  
 agcaactcaca aggaagcaca catttgcaaa tagcacctac cggaacagct ttgaatgtgc 420  
 cagcacagca ttgaacaggt tcgcgccctt actcctgtgc tctgttttct cgatcggaat 480  
 gttcgaaagt tgaaaagcgc attttttcat ctctcttttt ctattcttct tctgtttttt 540  
 atccctctct cgtcgtgttt tttctaaaca ttaccatact tcttcgccta cgaactcgcc 600  
 aagaaccaga acgcagcgtg cgtgcggtgc ttgcggtgtg tgtgtgtgtg tgtgtattcc 660  
 acggtgcgga gaagcaagat cggagaacag gcatcattcc cctttcacag acaattgcac 720  
 ttttgtacta gaacagaaaa cgagacagca taatttccaa cagcctcatt cactcatacc 780  
 aggctcacac cgacttttaa ccgaaacatg tactacagaa acaaaaacaa acaatatgga 840  
 gagtgctcgc gctgatacta agttaatatg aagagattac tggcgaggtc atcgatccca 900  
 tcccagacatc atcgctccag gctccagacc taccaagtcg cctaccatta cctaccacc 960  
 accgaccact actcacacag cattatcact tccgcgcgcg tcgcccgcgc cgccgacgcc 1020

gccgacgcca	ccacottcac	accgccctgc	caaaatgaat	gcgcattggt	gcgatagatt	1080
gaatttcoct	ggttgttggt	gttgttggtt	ttcttttgac	atgttttgtt	gttgtttttt	1140
ctttctctct	ctctctttct	gtggttccaa	catttcagac	gcatttttta	caccatatat	1200
aggtcacggt	gagtccttac	gaattataga	tgccgagttg	ggcacgctcg	agcatgtcca	1260
cagtggagca	acgccgcggc	gacgcggcct	gacgaggcgc	gagtcaaaact	cgggtaagta	1320
cgcgattgga	agtgggggga	cgtttaccct	accgtgtact	actacaacgc	actttacccc	1380
cacgcacacg	caccggcaga	cgcgaacgac	aacgatccgc	tggtggtcaa	cacggataag	1440
gggcgcaccc	gcggcattac	ggtcgatgcg	cccagcgga	agaagggtga	cgtgtggctc	1500
ggcattcoct	acgcccagcc	gccggtcggg	ccgctacggt	tccgtcatcc	cgggccggcc	1560
gaaaagtgga	ccggcgtgct	gaacacgacc	acaccgccca	acagctgcgt	gcagatcggt	1620
gacaccgtgt	tcggcgactt	cccgggcgcg	accatgtgga	acccgaacac	gcccctgtcc	1680
gaggactgtc	tgtacattaa	cgtggtggca	ccgcgacccc	ggcccaagaa	tgcggccgtc	1740
atgctgtgga	tcttcggcgg	cggcttctac	tccggcaccg	ccaccctgga	cgtgtacgac	1800
caccgggcgc	ttgcgtcgga	ggagaacgtg	atcgtggtgt	cgtgcagta	ccgcgtggcc	1860
agtctgggct	tcctgtttct	cggcaccccc	gaagcgccgg	gcaatgcggg	actgttcgat	1920
cagaaccttg	cgctacggta	ggtgtctttg	catgtgtgaa	tgagggtata	gtattctaac	1980
gaggtgctct	tcttcccac	acttcttggg	agtcagctgg	gtgcgggaca	acattcaccg	2040
gttcggtggc	gatccgtcgc	gtgtgacact	gttcggcgag	agtgccgggtg	ccgtctcggt	2100
gtcgtgcat	ctgctgtccg	ccctttcccg	cgtctgttcc	cagcggggcca	tcctgcagag	2160
cggctcgccg	acggcacccg	gggcattggt	atcgcgcgag	gaagccacac	taaggtacgt	2220
gccagctgct	gctttcccca	aaccaccaac	ccgcaacagc	tcacacaacc	ctcttttccg	2280
tcgctctttt	ctcgtccag	agcactgcgg	ttggccgagg	cggctcggtg	cccgcacgaa	2340
ccgagcaagc	tgagcgatgc	ggtcgagtgc	ctgcgcggca	aggacccgca	cgtgctgggt	2400
aacaacgagt	ggggcacgct	cggcatttgc	gagttcccgt	tcgtgccggt	ggtcgacggt	2460
gcgttcoctg	acgagacgcc	gcagcgttcg	ctcgccagcg	ggcgcttcaa	gaagacggag	2520
atcctcaccg	gcagcaacac	ggaggagggc	tactacttca	tcactacta	cctgaccgag	2580
ctgctgcgca	aggaggaggg	cgtgaccgtg	acgcgcgagg	agttcctgca	ggcggtgcgc	2640
gagctcaaac	cgtacgtgaa	cggggcgccg	cggcaggcga	tcgtgttcga	gtacaccgac	2700
tggaaccgagc	cggacaaccc	gaacagcaac	cgggacgcgc	tggaacaagat	ggtggggcgac	2760
tatcaacttca	cctgcaacgt	gaacgagttc	gcgcagcggg	acgccgagga	gggcaacaac	2820
gtctacatgt	atctgtacac	gcaccgcagc	aaaggcaacc	cgtggccgcg	ctggacgggc	2880
gtgatgcacg	gcgacgagat	caactacgtg	ttcggcgaac	cgtcaaccc	caccctcggc	2940
tacaccgagg	acgagaaaga	ctttagccgg	aagatcatgc	gatactgggt	caactttgcc	3000
aaaaccgggt	aagtgtgtgt	gtcaaacagc	agagtgtcga	tcgtctaac	accagcgtct	3060
tctctcttct	acagcaatcc	aaatcccaac	acggccagca	gcgaattccc	cgagtggccc	3120
aagcacaccc	cccacggacg	gcactatctg	gagctgggcc	tcaacacgtc	cttcgtcggt	3180
cggggcccac	ggttgaggca	gtgtgccttc	tggaagaagt	accttcccca	gctagtgtga	3240
gctacctgta	agtctcgtgc	agcacttgaa	acccctccc	acatcccat	cagggtccag	3300
gttgcaataa	taaatttcac	tttctctctc	tcacgtctct	tttcccaaaa	acagcgaacc	3360
taccagggcc	agcaccgcct	agtgaaccgt	gcgaaagcag	cgcatttttt	taccgacctg	3420
atctgatcgt	gctgctgggt	tcgctgctta	cggcgaccgt	cagattcata	caataattac	3480
taccccatcc	atggcctagt	tcgtttaagc	tttaagatag	tgaggaacaa	atttttccca	3540
aacaattttc	ccccctttag	agcagaaccg	aggagagat	aggactacat	agcgaanaag	3600
gaaaacaagt	ggtggcggac	gaggagagaa	gaagcaaatc	gaataatcga	agcaacaaca	3660
acaacaacaa	aaaaactgca	accgggttca	ctaaacccag	ggggcagctc	agtagcaaac	3720
tactacttaa	ataactactt	tcttatggca	aattatggca	agagcagtcg	tgatgggttc	3780
gatcagtatc	catctgaccg	gagcagctga	accgtttcat	gggcagttgc	tgcaatacac	3840
cacgaccggt	acacacagta	acacactttt	tatagcttta	cactaacaac	cactctcccc	3900
acgtctctct	tccccttccc	ctccacacag	acagcagcgc	cgtttgtagc	aggatctact	3960
accgtgcggg	ttgggtatgg	ggccaacaac	actaaacacc	acacatctac	taaaacacac	4020
cggaaacaata	aacaaatggt	aaacttacta	tatgaatata	catctagacg	catatatacg	4080
catgaactac	tacttcccct	cgtggctctga	caaaaacaca	ttacottgtc	cccccttccc	4140
cctccgggtt	gcttaccacc	actgaccccc	agtatgaatt	tgttccataa	taacgcttcg	4200
taactcgct						4209

&lt;210&gt; 23

&lt;211&gt; 2557

&lt;212&gt; ADN

&lt;213&gt; Anopheles gambiae souche KISUMU

&lt;400&gt; 23

```

aatgaatgcg cattgttgcg atagattgaa tttccttggg tgttgttggg gttgggttttc 60
ttttgacatg tttgtgtgtt gttttttctt tctctctctc tctctctgtg gttccaacat 120
ttcagacgca ttttttacac catatatagg tcacgggtgag tccgtacgaa ttatagatgc 180
cgagttgggc acgctcgagc atgtccacag tggagcaacg ccgcggcgac gcggtctgac 240
gaggcgcgag tccaactcgg gtaagtacgc gattggaagt ggggggacgt ttaccctgcc 300
gtgtactaca atgcacttta cccccaegca caegcaccgg cagacgcgaa cgacaacgat 360
ccgctgggtg tcaacacgga taagggggcg atccgcggca ttacggtcga tgcgcccagc 420
ggcaagaagg tggacgtgtg gctcggcatt ccctacgccc agccgcgggt cgggcccgtta 480
cggttccgtc atccgcggcc ggccgaaaag tggaccggcg tgctgaacac gaccacaccg 540
cccaacagct gcgtgcagat cgtggacacc gtgttcggcg acttcccggg cgcgaccatg 600
tggaaaccga acacgcccct gtccgaggac tgtctgtaca ttaacgtggg ggcaccgcga 660
ccccggccca agaatgcggc cgtcatgctg tggatcttcg gcggcggcct ctactccggc 720
accgccacc tggacgtgta cgaccaccgg gcgcttgctg cggaggagaa cgtgatcgtg 780
gtgtcgtgct agtaccgctg ggccagtgct ggcttcctgt ttctcggcac ccggaagcg 840
ccgggcaatg cgggactggt cgatcagaac ctgtcgctac ggtaggtgtc tttgcatggg 900
tgaatgaggg tatagtattc taacgaggtg ctcttcttcc catcacttct tgggagtcag 960
ctgggtgctg gacaacattc accggttcgg tggatgctg tcgctgtgta cactgttcgg 1020
cgagagtgcc ggtgccgtct cgggtgctgt gcatctgtgt tccgccctgt cccgcgatct 1080
gttccagcgg gccatcctgc agagcggctc gccgacggca ccgtgggcat tggtatcgcg 1140
cgaggaagcc acgctaaggt acgtgccagc tgctgctttc cccaaaccac caaccgcga 1200
cagctcacac aaccctcttt tccttcgctc tttctcgtc ccagagcact gcggttggcc 1260
gaggcggtcg gctgcccgca cgaaccgagc aagctgagcg atgcggtcga gtgtctgcgc 1320
ggcaaggatc cgcacgtgct ggtcaacaac gagtggggca cgctcggcat ttgcgagttc 1380
ccgttcgtgc cgggtggtcga cgggtgcgtt ctggacgaga ccgcgcagcg ttcgctcgcc 1440
agcgggctgc tcaagaagac ggagatcttc accggcagca acacggagga gggctactac 1500
ttcatcatct actacctgac cgagctgctg cgcaaggagg agggcggtgac cgtgacgcgc 1560
gaggagtcc tgcaggcggg gcgcgagctc aaccgcgtac tgaacggggc ggcccggcag 1620
gcgatcgtgt tcgagtacac cgactggacc gagccggaca acccgaacag caaccgggac 1680
gcgctggaca agatggtggg cgactatcac ttcacctgca acgtgaacga gttcgcgag 1740
cggtagcgcg aggagggcaa caacgtctac atgtatctgt acacgcaccg cagcaaaggc 1800
aaccgctggc cgcgctggac gggcgtgatg cagggcgacg agatcaacta cgtgttcggc 1860
gaaccgctca accccaccct cggctacacc gaggacgaga aagactttag ccggaagatc 1920
atgcgatact ggtctaactt tgccaaaacc ggttaagtgt gtgtgtgtgt gtgtgtcaaa 1980
cagcagagtg tcatcgctc taacgccttc tctcttcaac agcaatccaa atcccaacac 2040
ggccagcagc gaattccccg agtggcccaa gcacaccgcc cacggacggc actatctgga 2100
gctgggcctc aacacgtcct tcgtcggtcg gggcccacgg ttgaggcagt gtgccttctg 2160
gaagaagtac cttccccagc tagttgcagc tacctgtaag tctcgtgcag cgcttgaaat 2220
cctctccgc atcctcaaca ggggccaggg tgcaataaca aatgtatctc tctctctctc 2280
acgtctcttt tccccaaaac agcgaacctc ccaggggccag caccgcccag tgaaccgtgc 2340
gaaagcagcg cattttttta ccgacctgat ctgatcgtgc tgctgggtgtc gctgcttacg 2400
gcgaccgtca gattcatata ataattacta cccatccat ggccatagttc ttttaagctt 2460
taagatagtg aggaacaaat ttttcctaac caatttccca acccccttta gagcagaacc 2520
gaggagagaga taggactaca tagcgaaaag ggaaaac 2557

```

&lt;210&gt; 24

&lt;211&gt; 273

&lt;212&gt; ADN

&lt;213&gt; Culex pipiens souche S-LAB

&lt;400&gt; 24

```

attgaaccgg acaaccggaa cagcaaccgt gacgcgctgg acaagatggg cggggattat 60
cacttcacct gcaacgtgaa cgaattcgcc cagcggtagc ccgaggaggg caacaacgtg 120
ttcatgtacc tgtacacgca cagaagcaaa ggaaatccct ggccgaggtg gaccggcggtg 180

```

atgcacggcg acgagatcaa ctacgtgttt ggcgaaccgc tgaactcggc cctcgggtac 240  
caggacgacg agaaggactt tagccggaaa att 273

<210> 25

<211> 273

<212> ADN

<213> Culex pipiens souche SR

<400> 25

atcgaaccgg acaaccggaa cagcaaccgt gacgcgctcg acaagatggg cggggattat 60  
cacttcacct gcaacgtgaa cgagttcgcc cagcggtagc ccgaggaggg caacaatgtg 120  
ttcatgtacc tgtacacgca cagaagcaaa ggaaatccct ggccgaggtg gactggcgtg 180  
atgcacggcg acgagatcaa ctacgtgttt ggcgaaccgc tgaactcggc cctcgggtac 240  
caggacgacg agaaggactt tagccggaaa att 273

<210> 26

<211> 273

<212> ADN

<213> Aedes aegypti

<400> 26

actgaaccgg aaaatcccaa cagcaatcgg gatgcattgg acaaaatggg cggagattat 60  
cacttcacgt gtaatgtgaa tgagtttgcc cagcgatatg cagaagaagg caacaatgtg 120  
tacatgtatc tgtacactca tagaagcaaa ggtaaccctt ggccacgggtg gaccgggtgtg 180  
atgcatgggtg acgagatcaa ttatgtgttc ggtgagcctc tgaactctga tctgggggtac 240  
atggaggtatg aaaaagactt cagtaggaag att 273

<210> 27

<211> 273

<212> ADN

<213> Aedes albopictus

<400> 27

actgaaccag agaatcccaa cagcaatcgg gatgcgttgg acaaaatggg gggagattat 60  
catttcacct gcaacgtgaa cgagtttgcc cagcgatatg cggaagaggg caacaacgtg 120  
tacatgtatt tgtacactca cagaagcaaa ggtaaccctt ggccacgggtg gaccgggggtg 180  
atgcatgggtg acgagatcaa ctatgtattc ggtgagccgt tgaattccga cctgggggtac 240  
atggacgatg agaaagattt cagtagaaag ata 273

<210> 28

<211> 273

<212> ADN

<213> Anopheles darlingi

<400> 28

acagaaccgg acaaccggaa cagtaaccgg gacgcgctgg acaagatggg cgggtgattat 60  
cacttcacgt gtaacgtcaa tgagtttgcg cagcggtagc ccgaggaggg caacaacgtc 120  
tacatgtatc tgtacacgca ccgtagcaaa ggcaaccctt ggccccgctg gaccgggggtg 180  
atgcatgggtg atgagattaa ctacgtgttc ggtgaaccgc tcaaccggac gctcgggttac 240  
accgacgatg agaagggttt cagccgggaag att 273

<210> 29

<211> 273

&lt;212&gt; ADN

<213> *Anopheles sundaicus*

&lt;400&gt; 29

accgagccgg acaacccgaa cagcaaccga gacgcgctgg acaagatggt cggcgactat 60  
 cacttcacct gcaacgtcaa cgagttcgcc cagcgggtacg ccgaggaggg caacaacgtc 120  
 tacatgtatc tgtacacgca ccgaagcaaa ggcaaccggt ggccacgctg gacgggtgtg 180  
 atgcacgggtg acgagattaa ttacgtgttt ggagagccgc ttaacccac gtcgggatac 240  
 accgaggacg agaagactt tagccggaag atc 273

&lt;210&gt; 30

&lt;211&gt; 273

&lt;212&gt; ADN

<213> *Anopheles minimus*

&lt;400&gt; 30

accgaaccag ataatccgaa cagcaaccgg gacgcactgg acaagatggt gggcgactac 60  
 catttcacct gtaacgtgaa cgagttcgca cagcgggtacg ccgaggaggg caacaatgta 120  
 tacatgtacc tgtacacgca ccgaagcaaa ggcaaccggt ggccacgctg gaccggcggt 180  
 atgcacgggtg acgagattaa ctacgtgttc ggggaaccgc tcaaccaag cctcgggctac 240  
 accgaagacg agaaagactt tagccggaag atc 273

&lt;210&gt; 31

&lt;211&gt; 273

&lt;212&gt; ADN

<213> *Anopheles moucheti*

&lt;400&gt; 31

accgaaccag ataatccgaa cagcaaccgg gacgcactgg acaagatggt gggcgactac 60  
 catttcacct gtaacgtgaa cgagttcgca cagcgggtacg ccgaggaggg caacaatgta 120  
 tacatgtacc tgtacacgca ccgaagcaaa ggcaaccggt ggccacgctg gaccggcggt 180  
 atgcacgggtg acgagattaa ctacgtgttc ggggaaccgc tcaaccaag cctcgggctac 240  
 accgaagacg agaaagactt tagccggaag atc 273

&lt;210&gt; 32

&lt;211&gt; 273

&lt;212&gt; ADN

<213> *Anopheles arabiensis*

&lt;400&gt; 32

accgagccgg acaacccgaa cagcaaccgg gacgcggttg acaagatggt gggcgactat 60  
 cacttcacct gcaacgtgaa cgagttcgcg cagcgggtacg ccgaggaggg caacaacgtc 120  
 tacatgtatc tgtacacgca ccgcagcaaa ggcaaccggt ggccgcgctg gacgggctg 180  
 atgcacggcg acgagatcaa ctacgtgttc ggcgaaccgc tcaacccac cctcgggctac 240  
 accgaggacg agaaagactt tagccggaag atc 273

&lt;210&gt; 33

&lt;211&gt; 273

&lt;212&gt; ADN

<213> *Anopheles funestus*

&lt;400&gt; 33

accgagccgg acaacccgaa cagcaaccgt gacgcgctcg acaaaatggt gggcgactat 60  
 catttcacct gcaacgtgaa cgagttcgcc cagcgggtacg ccgaggaggg caacaatgta 120



tacatgtacc tgtacacgca ccgaagcaaa ggcaacccat ggccacgctg gacgggctt 180  
 atgcacggtg atgagattaa ctatgtgttc ggggaaccgc tcaatcccag cctcggctac 240  
 accgaggacg agaaagactt tagccggaag atc 273

<210> 34

<211> 273

<212> ADN

<213> *Anopheles pseudopunctipennis*

<400> 34

accgagccgg acaacccgaa cagcaaccgg gacgcgctgg acaagatggt gggcgactac 60  
 cacttcacgt gcaacgtgaa cgagttcgcc cagcgctacg ccgaagaggg caacaacgtg 120  
 tacatgtatc tgtacacgca ccgaagcaaa ggcaacccgt ggccgcgctg gaccggcgctc 180  
 atgcatgggg acgagattaa ctacgtgttt ggggaaccgc ttaaccggg gctcggctat 240  
 accgaggacg agaaggactt tagccgcaag atc 273

<210> 35

<211> 273

<212> ADN

<213> *Anopheles sacharovi*

<400> 35

accgagccgg acaacccgaa cagcaaccgg gacgcgctgg acaagatggt cggcgactac 60  
 cacttcacct gcaacgtgaa cgagttcgcg cagcggtacg ccgaggaggg caacaacgtc 120  
 tacatgtacc tgtacacgca caggagcaaa ggcaacccat ggccgcgctg gaccggcgctc 180  
 atgcatggcg acgagatcaa ctacgtgttc ggcgaaccgc tcaatcccag cctaggctac 240  
 accgatgacg agaaagactt tagccggaag att 273

<210> 36

<211> 273

<212> ADN

<213> *Anopheles stephensi*

<400> 36

accgaaccgg acaatccgaa cagcaaccgg gatgcattgg acaaaatggt gggcgattac 60  
 catttcacgt gcaacgtgaa cgagttcgca cagcgatacg ccgaggaggg caacaatgtg 120  
 tacatgtatc tgtacacgca ccgaagcaaa ggcaatccgt ggccacgctg gaccggcgctt 180  
 atgcatgggg acgaaattaa ctacgtgttc ggggaaccgc tcaaccctag ccttggttac 240  
 accgacgacg agaaagactt tagccggaag atc 273

<210> 37

<211> 273

<212> ADN

<213> *Anopheles albimanus*

<400> 37

acggagccgg acaatccgaa cagcaaccgg gacgcactgg acaagatggt cggcgattat 60  
 cactttacgt gcaacgtcaa cgagttcgcg cagcggtacg ccgaggaggg caacaacgtc 120  
 tacatgtatc tgtatagca cgcagcaaa ggcaatccgt ggccccgttg gacgggctg 180  
 atgcatggcg atgagatcaa ctacgtgttt ggtgaaccgc tgaaccgcac gctcggctac 240  
 accgacgacg agaagggtt cagccggaag atc 273

<210> 38

<211> 273  
 <212> ADN  
 <213> Anopheles nili

<400> 38  
 accgagccgg ataacccgaa cagcaaccgg gacgcgttag acaagatggt gggcgactac 60  
 cacttcacgt gcaacgtgaa cgagttcgcc cagcggtagc ccgaggaggg caacaacgtc 120  
 tacatgtacc tctacacgca ccggagcaaa ggcaatccct ggccgcgttg gacgggctc 180  
 atgcacggtg acgagatcaa ctacgtgttc ggggaaccgc ttaacccgag cctcggctac 240  
 accgaggacg-agaaggaett-cagccycaag atg 273

<210> 39  
 <211> 24  
 <212> ADN  
 <213> Séquence artificielle

<400> 39  
 atmgwgattyg agtacacsga ytg 24

<210> 40  
 <211> 24  
 <212> ADN  
 <213> Séquence artificielle

<400> 40  
 ggcaaarttk gwccagtatc kcat 24

<210> 41  
 <211> 20  
 <212> ADN  
 <213> Séquence artificielle

<220>  
 <223> Description de la séquence artificielle:amorce

<400> 41  
 ggygckacma tgtggaaycc 20

<210> 42  
 <211> 24  
 <212> ADN  
 <213> Séquence artificielle

<220>  
 <223> Description de la séquence artificielle:amorce

<400> 42  
 accamratca cgattyctc cgac 24

<210> 43  
 <211> 23  
 <212> ADN  
 <213> Séquence artificielle

<220>

<223> Description de la séquence artificielle:amorce

<400> 43

tacatcaacg tggtcgtgcc acg

23

<210> 44

<211> 19

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle:amorce

<400> 44

gtcacggttg ctgttcggg

19

<210> 45

<211> 16

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle:amorce

<400> 45

cgacgccacc ttcaca

16

<210> 46

<211> 20

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle:amorce

<400> 46

gatggcccgc tggaacagat

20

<210> 47

<211> 20

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle:amorce

<400> 47

gggtgcggga caacattcac

20

<210> 48

<211> 17

&lt;212&gt; ADN

&lt;213&gt; Séquence artificielle

&lt;220&gt;

&lt;223&gt; Description de la séquence artificielle:amorce

&lt;400&gt; 48

ccccgaccga cgaagga

17

&lt;210&gt; 49

&lt;211&gt; 20

&lt;212&gt; ADN

&lt;213&gt; Séquence artificielle

&lt;220&gt;

&lt;223&gt; Description de la séquence artificielle:amorce

&lt;400&gt; 49

agatggtggg cgactatcac

20

&lt;210&gt; 50

&lt;211&gt; 20

&lt;212&gt; ADN

&lt;213&gt; Séquence artificielle

&lt;220&gt;

&lt;223&gt; Description de la séquence artificielle:amorce

&lt;400&gt; 50

ctcgtccgcc accacttggt

20

&lt;210&gt; 51

&lt;211&gt; 585

&lt;212&gt; PRT

&lt;213&gt; Ciona intestinalis

&lt;400&gt; 51

Leu	Pro	Arg	Tyr	Gly	Ser	Val	Arg	Gly	Lys	His	Val	Glu	Ser	Pro	Pro
1				5					10					15	

Arg	His	Gln	Arg	Ile	Ala	Ala	Phe	Leu	Gly	Ile	Pro	Phe	Ala	Ser	Pro
		20						25				30			

Pro	Val	Gly	Glu	Leu	Arg	Phe	Ala	Ala	Pro	Gln	Pro	Pro	Leu	Ser	Trp
		35					40					45			

Glu	Pro	Asp	Val	Arg	Gln	Thr	Thr	Glu	Phe	Gly	Asn	Ser	Cys	Val	Gln
	50					55					60				

Ile	Asp	Asp	Glu	Val	Phe	Gly	Asn	Phe	Arg	Glu	Met	Trp	Asn	Ala	Pro
65					70					75				80	

Asn	Leu	Lys	Ser	Glu	Asp	Cys	Leu	Tyr	Leu	Asn	Ile	Trp	Thr	Pro	Arg
				85					90					95	

Ile Pro Thr Ser Thr Arg Ser Gln Pro Leu Ala Val Met Val Trp Ile  
 100 105 110  
 Tyr Gly Gly Ser Phe Tyr Ser Gly Thr Thr Ala Leu Ala Leu Tyr Asp  
 115 120 125  
 Gly Arg Tyr Leu Ala Ala Gln Gly Gly Val Val Val Val Ser Ile Asn  
 130 135 140  
 Tyr Arg Leu Gly Pro Leu Gly Phe Leu Ala Pro Leu Ala Gly Thr Pro  
 145 150 155 160  
 Gly Asn Ala Gly Leu Leu Asp Gln Gln Leu Ala Leu Lys Trp Val Arg  
 165 170 175  
 Asp Asn Ile Arg Ala Phe Gly Gly Asn Pro Asp Asn Val Thr Leu Met  
 180 185 190  
 Gly Glu Ser Ala Gly Ala Ala Ser Ile Gly Leu His Thr Val Ala Pro  
 195 200 205  
 Ser Ser Arg Gly Leu Phe Asn Arg Val Ile Phe Gln Ser Gly Asn Gln  
 210 215 220  
 Met Thr Pro Trp Ser Thr Ile Ser Leu Pro Thr Ser Leu Asn Arg Thr  
 225 230 235 240  
 Arg Ile Leu Ala Ala Asn Leu Arg Cys Pro Asn Pro Arg Thr Ser Ser  
 245 250 255  
 Glu Leu Asp Val Leu Thr Cys Leu Arg Ser His Ser Ala Val Asp Val  
 260 265 270  
 Phe Ser Asn Ser Trp Ile Thr Gln Glu Ile Phe Asp Phe Pro Phe Val  
 275 280 285  
 Pro Val His Gly Thr Ser Phe Leu Pro Glu His Pro His Glu Val Thr  
 290 295 300  
 Arg Lys Gly Glu Gln Ala Asp Val Asp Val Met Ala Gly His Asn Thr  
 305 310 315 320  
 Asn Glu Gly Ser Tyr Phe Thr Leu Tyr Thr Val Pro Gly Phe Asn Ile  
 325 330 335  
 Ser Ser Gln Ser Ile Leu Ser Lys Lys Glu Tyr Ile Asp Gly Ile Ala  
 340 345 350  
 Leu Ser Gly Ile Lys Thr Asn Glu Leu Gly Arg Ser Gly Ala Ala Phe  
 355 360 365  
 Met Tyr Ala Asp Trp Glu Asn Pro Asp Asn Val Leu Gln Tyr Arg Asp  
 370 375 380  
 Gly Val Asn Glu Ile Val Gly Asp Phe His Val Val Cys Pro Thr Val  
 385 390 395 400

Leu Leu Thr Lys Arg His Ser Arg Thr Phe Ser Asn Asn Asn Val Tyr  
405 410 415

Leu Tyr His Leu Ser Tyr Arg Leu Ser Asn Asn Pro Trp Pro Ala Trp  
420 425 430

Met Gly Val Met His Gly Tyr Glu Ile Glu Leu Met Phe Gly Thr Pro  
435 440 445

---

~~Trp Phe Gly Thr Ser Gln Phe Thr Ser Gly Tyr Asn Asp Val Asp Arg~~  
450 455 460

Ser Val Ser Arg Arg Met Val His Tyr Trp Thr Asn Phe Ala Lys Phe  
465 470 475 480

Gly Asn Pro Asn Gly Leu Arg Ser Ala Asn Glu Leu Asp Leu Arg Ser  
485 490 495

Thr Asp Trp Pro Arg Phe Asp Asp Val Arg Gln Arg Tyr Leu Glu Ile  
500 505 510

Gly Ile Asp Asp Asp Val Met Gly Pro Phe Pro Asn Ser Phe Arg Cys  
515 520 525

Ala Phe Trp Glu Arg Tyr Leu Pro Ser Leu Lys Leu Ala Ser Ser Ala  
530 535 540

Asp Met Asp Glu Val Glu Thr Lys Trp Lys Ile Glu Phe Asn Arg Trp  
545 550 555 560

Thr Glu Ser Met Asp Leu Trp Asp Arg Ser Phe Lys Ala Tyr Ser Lys  
565 570 575

Asp Gly Lys Gln Ser Ser Cys Pro Asn  
580 585

<210> 52

<211> 583

<212> PRT

<213> Ciona savignyi

<400> 52

Gly Ser Ile Gln Gly Lys His Val Glu Val Thr Ala His Arg Gln Arg  
1 5 10 15

Tyr Gly Arg Val Ala Thr Phe Gln Gly Ile Pro Phe Ala Gln Pro Pro  
20 25 30

Val Gly Glu Leu Arg Phe Ala Ala Pro Gln Pro Pro Leu Ser Trp Glu  
35 40 45

Pro Asp Val Lys Met Thr Ser Glu Phe Gly Asn Ser Cys Ile Gln Glu  
50 55 60

Asp Asp Leu Val Phe Gly Asn Phe Thr Gly Gly Ser Gln Met Trp Asn  
65 70 75 80

Ser Pro Asn Ala Lys Ser Glu Asp Cys Leu Tyr Leu Asn Val Trp Thr  
 85 90 95  
 Pro Val Arg Ser Arg His Ala Glu Pro Leu Ala Val Leu Val Trp Ile  
 100 105 110  
 Tyr Gly Gly Ser Tyr Tyr Ser Gly Thr Ser Ser Leu Ala Leu Tyr Asp  
 115 120 125  
 Gly Arg Tyr Leu Ala Ala Thr Gly Gly Val Val Val Val Ser Leu Asn  
 130 135 140  
 Tyr Arg Leu Gly Pro Ile Gly Phe Leu Ala Pro Leu Ala Asp Glu Thr  
 145 150 155 160  
 Pro Gly Asn Val Gly Leu Leu Asp Gln Gln Leu Ala Leu Lys Trp Val  
 165 170 175  
 Arg Asp Asn Ile Arg Glu Phe Gly Gly Asn Pro Asn Asn Val Thr Val  
 180 185 190  
 Met Gly Glu Ser Ala Gly Ala Ala Ser Ile Gly Leu His Thr Ile Ala  
 195 200 205  
 Pro Ser Ser Arg Gly Leu Phe Ser Arg Val Ile Leu Gln Ser Gly Asn  
 210 215 220  
 Gln Met Thr Pro Trp Ser Thr Ile Ser Leu Glu Thr Ser Leu Asn Arg  
 225 230 235 240  
 Thr Arg Thr Leu Ala Ala Asn Leu Asn Cys Pro Lys Pro Arg Thr Ala  
 245 250 255  
 Ser Glu Ala Asp Ile Leu Ala Cys Leu Arg Thr His Thr Ala Asn Glu  
 260 265 270  
 Val Phe Ala Gly Ser Trp Ile Thr Lys Glu Ile Phe Asp Phe Pro Phe  
 275 280 285  
 Val Pro Val His Gly Thr Thr Phe Leu Pro Glu His Pro His Glu Val  
 290 295 300  
 Thr Arg Arg Gly Asp Gln Ala Glu Val Asp Val Leu Ala Gly Tyr Asn  
 305 310 315 320  
 Thr Asn Glu Gly Ser Tyr Phe Thr Ile Tyr Thr Val Pro Gly Tyr Asn  
 325 330 335  
 Ile Thr Thr Asn Ser Val Leu Asn Arg Arg Gln Tyr Leu Ala Gly Val  
 340 345 350  
 Asp Leu Ser Gly Leu Lys Thr Asn Thr Met Gly Arg Ser Ala Ala Ala  
 355 360 365  
 Phe Met Tyr Thr Asp Trp Glu Asn Leu Asp Asn Glu Leu Gln Tyr Arg  
 370 375 380

Asp Ala Val Asn Glu Ile Val Gly Asp Phe His Val Val Cys Pro Thr  
385 390 395 400

Val Leu Val Ser Lys Arg His Ser Asn Ser Phe Pro Asn Arg Asn Val  
405 410 415

Phe Leu Tyr His Leu Ser Tyr Arg Val Ser Thr Asn Pro Trp Pro Ile  
420 425 430

Trp Met Gly Val Met His Gly Tyr Glu Ile Glu Leu Met Phe Gly Thr  
435 440 445

Pro Trp Phe Gly Asn Ser Lys Phe Thr Arg Gly Tyr Ser Asp Leu Asp  
450 455 460

Arg Ser Val Ser Arg Arg Met Val Arg Tyr Trp Thr Asn Phe Ala Lys  
465 470 475 480

Phe Gly Asn Pro Asn Gly Leu Arg Asn Gln Asn Gln Glu Leu Val Ser  
485 490 495

Asp Trp Pro Arg Phe Asn Asp Val Thr Gln Arg Tyr Leu Glu Ile Ala  
500 505 510

Asp Asp Asp Val Thr Met Ala Pro Phe Pro Asp Ser Phe Arg Cys Ala  
515 520 525

Phe Trp Gln Lys Tyr Leu Pro Ser Leu Gln Leu Ala Ser Ser Asn Met  
530 535 540

Asp Glu Val Glu Thr Lys Trp Lys Ile Glu Phe His Arg Trp Ser Glu  
545 550 555 560

Ser Met Asp Leu Trp Asp Arg Ser Phe Lys Ala Tyr Ser Ser Asp Asp  
565 570 575

Lys Gln Asn Ser Cys Pro Asn  
580

<210> 53

<211> 645

<212> PRT

<213> Anopheles gambiae

<400> 53

Met Ala Ser Ala Tyr Tyr His Gln Ser Ala Val Gly Val Gly Asn Val  
1 5 10 15

Leu Val Leu Leu Leu Gly Ala Thr Val Ile Cys Pro Ala Tyr Ala Ile  
20 25 30

Ile Asp Arg Leu Val Val Gln Thr Ser Ser Gly Pro Ile Arg Gly Arg  
35 40 45

Ser Thr Met Val Gln Gly Arg Glu Val His Val Phe Asn Gly Val Pro  
50 55 60



Phe Ala Lys Pro Pro Val Asp Ser Leu Arg Phe Lys Lys Pro Val Pro  
 65 70 75 80  
 Ala Glu Pro Trp His Gly Val Leu Asp Ala Thr Arg Leu Pro Pro Ser  
 85 90 95  
 Cys Ile Gln Glu Arg Tyr Glu Tyr Phe Pro Gly Phe Ala Gly Glu Glu  
 100 105 110  
 Met Trp Asn Pro Asn Thr Asn Val Ser Glu Asp Cys Leu Tyr Leu Asn  
 115 120 125  
 Ile Trp Val Pro Thr Lys Thr Arg Leu Arg His Gly Arg Gly Leu Asn  
 130 135 140  
 Phe Gly Ser Asn Asp Tyr Phe Gln Asp Asp Asp Phe Gln Arg Gln  
 145 150 155 160  
 His Gln Ser Lys Gly Gly Leu Ala Met Leu Val Trp Ile Tyr Gly Gly  
 165 170 175  
 Gly Phe Met Ser Gly Thr Ser Thr Leu Asp Ile Tyr Asn Ala Glu Ile  
 180 185 190  
 Leu Ala Ala Val Gly Asn Val Ile Val Ala Ser Met Gln Tyr Arg Val  
 195 200 205  
 Gly Ala Phe Gly Phe Leu Tyr Leu Ala Pro Tyr Ile Asn Gly Tyr Glu  
 210 215 220  
 Glu Asp Ala Pro Gly Asn Met Gly Met Trp Asp Gln Ala Leu Ala Ile  
 225 230 235 240  
 Arg Trp Leu Lys Glu Asn Ala Lys Ala Phe Gly Gly Asp Pro Asp Leu  
 245 250 255  
 Ile Thr Leu Phe Gly Glu Ser Ala Gly Gly Ser Ser Val Ser Leu His  
 260 265 270  
 Leu Leu Ser Pro Val Thr Arg Gly Leu Ser Lys Arg Gly Ile Leu Gln  
 275 280 285  
 Ser Gly Thr Leu Asn Ala Pro Trp Ser His Met Thr Ala Glu Lys Ala  
 290 295 300  
 Leu Gln Ile Ala Glu Gly Leu Ile Asp Asp Cys Asn Cys Asn Leu Thr  
 305 310 315 320  
 Met Leu Lys Glu Ser Pro Ser Thr Val Met Gln Cys Met Arg Asn Val  
 325 330 335  
 Asp Ala Lys Thr Ile Ser Val Gln Gln Trp Asn Ser Tyr Ser Gly Ile  
 340 345 350  
 Leu Gly Phe Pro Ser Ala Pro Thr Ile Asp Gly Val Phe Met Thr Ala  
 355 360 365

Asp Pro Met Thr Met Leu Arg Glu Ala Asn Leu Glu Gly Ile Asp Ile  
 370 375 380  
 Leu Val Gly Ser Asn Arg Asp Glu Gly Thr Tyr Phe Leu Leu Tyr Asp  
 385 390 395 400  
 Phe Ile Asp Tyr Phe Glu Lys Asp Ala Ala Thr Ser Leu Pro Arg Asp  
 405 410 415  


---

 Lys Phe Leu Glu Ile Met Asn Thr Ile Phe Asn Lys Ala Ser Glu Pro  
 420 425 430  
 Glu Arg Glu Ala Ile Ile Phe Gln Tyr Thr Gly Trp Glu Ser Gly Asn  
 435 440 445  
 Asp Gly Tyr Gln Asn Gln His Gln Val Gly Arg Ala Val Gly Asp His  
 450 455 460  
 Phe Phe Ile Cys Pro Thr Asn Glu Phe Ala Leu Gly Leu Thr Glu Arg  
 465 470 475 480  
 Gly Ala Ser Val His Tyr Tyr Tyr Phe Thr His Arg Thr Ser Thr Ser  
 485 490 495  
 Leu Trp Gly Glu Trp Met Gly Val Leu His Gly Asp Glu Val Glu Tyr  
 500 505 510  
 Ile Phe Gly Gln Pro Met Asn Ala Ser Leu Gln Tyr Arg Gln Arg Glu  
 515 520 525  
 Arg Asp Leu Ser Arg Arg Met Val Leu Ser Val Ser Glu Phe Ala Arg  
 530 535 540  
 Thr Gly Asn Pro Ala Leu Glu Gly Glu His Trp Pro Leu Tyr Thr Arg  
 545 550 555 560  
 Glu Asn Pro Ile Tyr Phe Ile Phe Asn Ala Glu Gly Glu Asp Asp Leu  
 565 570 575  
 Arg Gly Glu Lys Tyr Gly Arg Gly Pro Met Ala Thr Ser Cys Ala Phe  
 580 585 590  
 Trp Asn Asp Phe Leu Pro Arg Leu Arg Ala Trp Ser Val Pro Leu Lys  
 595 600 605  
 Asp Pro Cys Lys Leu Asp Asp His Thr Ser Ile Ala Ser Thr Ala Arg  
 610 615 620  
 Ala Ala Pro Thr Val Ala Leu Leu Ile Ala Leu Ser Leu Ala Val Ala  
 625 630 635 640  
 Arg Leu Val Ala Ala  
 645

This Page is inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images  
problems checked, please do not report the  
problems to the IFW Image Problem Mailbox**